


无人机航空遥感图像拼接技术研究

程远航 著

清华大学出版社

无人机航空遥感图像拼接技术研究

程远航　著

清华大学出版社
北　京

内 容 简 介

无人机航空遥感系统具有图像分辨率高、图像实时传输、适合高危地区作业、成本低、机动灵活等优点，适用于低空高分辨率遥感数据的实时获取，在区域性、工程性、灾害性和军事性的遥感监测中发挥着大型遥感系统难以替代的作用。无人机执行遥感监测任务时，需要实时传输其所获取的图像以及状态数据，这就要求无人机航空遥感系统具备自动、高速地完成图像的获取、压缩、传输、处理、显示以及存储等功能。其中，确保遥感图像处理的精准性、实时性与可视性是无人机得以有效利用的重要前提条件。由于现有成像设备的性能所限，目前的航空遥感成像系统无法获得大面积、高分辨率的观测图像，因此需要将获取的序列遥感图像进行在线动态拼接，以提高遥感图像的信息获取能力。本书结合无人机航空遥感具体的应用需求，针对航空遥感图像拼接技术中的几项关键技术进行了研究和探讨。内容包括：图像几何畸变校正、灰度序列遥感图像拼接、颜色空间与颜色相似性度量、彩色遥感图像拼接、图像特征提取方法、SVM 机器学习算法、可变形模板的多目标跟踪模型等。

本书可供信息科学、遥感科学和数理科学的科研人员和工程技术人员阅读，还可以作为计算机科学与技术、自动化、电子工程、遥感技术等专业的研究生和高年级本科生的教学参考书。

图书在版编目(CIP)数据

无人机航空遥感图像拼接技术研究 / 程远航 著. —北京：清华大学出版社，2016
ISBN 978-7-302-44403-9

Ⅰ. ①无… Ⅱ. ①程… Ⅲ. ①无人驾驶飞机—航空遥感—遥感图象—图象处理—研究 Ⅳ. ①TP72

中国版本图书馆 CIP 数据核字(2016)第 168724 号

责任编辑：施 猛 王旭阳
封面设计：常雪影
版式设计：方加青
责任校对：牛艳敏
责任印制：宋 林

出版发行：清华大学出版社
网 址：http://www.tup.com.cn，http://www.wqbook.com
地 址：北京清华大学学研大厦 A 座 邮 编：100084
社 总 机：010-62770175 邮 购：010-62786544
投稿与读者服务：010-62776969，c-service@tup.tsinghua.edu.cn
质 量 反 馈：010-62772015，zhiliang@tup.tsinghua.edu.cn
印 装 者：北京密云胶印厂
经 销：全国新华书店
开 本：185mm×230mm 印 张：12.25 字 数：216 千字
版 次：2016 年 8 月第 1 版 印 次：2016 年 8 月第 1 次印刷
印 数：1 ～ 1000
定 价：48.00 元

产品编号：070381-01

前言

无人机航空遥感系统具有图像分辨率高、图像实时传输、适合高危地区作业、成本低、机动灵活等优点，适用于低空高分辨率遥感数据的实时获取，在区域性、工程性、灾害性和军事性的遥感监测中发挥着大型遥感系统难以替代的作用。无人机执行遥感监测任务时，需要实时传输其所获取的图像以及状态数据，这就要求无人机航空遥感系统具备自动、高速地完成图像的获取、压缩、传输、处理、显示以及存储等功能。其中，确保遥感图像处理的精准性、实时性与可视性是无人机得以有效利用的重要前提条件。由于现有成像设备的性能所限，目前的航空遥感成像系统无法获得大面积、高分辨率的观测图像，因此需要将获取的序列遥感图像进行在线动态拼接，以提高遥感图像的信息获取能力。

结合无人机航空遥感具体的应用需求，本书针对航空遥感图像拼接技术进行了以下研究工作。

(1) 根据无人机遥感图像成像的内、外方位元素，采用直角空间变换及二次线性插补方法，实现了遥感图像校正。根据获取的无人机飞行状态参数，实现了连续两幅无人机航空遥感快视图像之间重叠区域的图像范围计算。

(2) 提出一种分布存储环境下的并行几何校正算法，每个处理器通过计算本地输入子图像在目标图像中的范围，确定其需要进行重采样计算的区域，使计算过程中所需的数据均为本地数据，很好地解决了数据局部性问题。利用首尾相连的闭线段近似表示理想的输出图像块边界，详细讨论了局部输出区域的计算方法，并采用一种新的存储结构用于保存校正后的输出图像块信息。

(3) 基于图像数据总体分布的统计，分析了飞行试验图像的成像质量。基于人眼亮度视觉特性曲线，结合小波变换和Curvelet变换特点，提出一种新的图像增强方法，实现了

无人机遥感序列图像的自适应增强处理。

(4) 提出结合小波变换和Canny算法的边缘提取算法，该算法将Canny算法的非极大值抑制、双阈值算法与小波变换多尺度分析思想相结合，利用Canny算法在高分辨率图像的精确定位性与小波变换后的低分辨率图像的抗噪声能力相结合，从而消除噪声的干扰，提取更加完整的边缘。

(5) 将图像匹配分为粗匹配和精匹配两个步骤。在粗匹配计算中，首先确定待拼接的两幅图像之间的大致重叠区域，然后利用小波变换与Canny算法提取图像边缘，最后采用区域匹配法求得两幅图像的匹配点。在精匹配计算中，首先确定更小的搜索区域，然后利用最小二乘法计算得到待拼接图像之间的最佳相对位置关系，最后根据匹配结果，实现两幅图像的拼接。

(6) 基于人眼的颜色视觉特性分析，本书提出了一种具有抗亮度干扰能力的彩色图像色差度量方法。利用颜色相似性分析，并引入协方差矩阵计算，给出了彩色图像特征模板的提取方法。利用最小二乘法，建立了两幅彩色图像之间的亮度变换函数曲线，实现了基于基准图像亮度分布的伽马校正。

(7) 目标的特征描述是跟踪算法中的重要步骤，直接影响最终的跟踪效果，决定整个跟踪系统的性能。基于分类学习的目标跟踪一般不选择较直观的整体特征，而偏向由基本特征进行局部特征统计，使特征描述子更易于区分且计算量较小。选择目前成功应用于目标检测的HOG特征作为目标特征。

(8) 传统的离线分类器需要人为标记大量的样本训练分类器，当待测样本与训练样本具有较大差异时，无法及时适应目标的外观变化，且难以推广到其他目标类型，分类器应用具有局限性。基于检测的跟踪算法，在跟踪过程中，跟踪器根据上一帧的跟踪结果信息在线更新分类器，然后再对当前帧检测，以获取当前运动目标的准确预测位置，有效抑制跟踪漂移。首先构造目标的外观模型。在初始帧中随机选择大量含目标区域以及含背景区域等的大图像块，分别提取HOG特征，标记为正负样本，然后将这些样本特征集合作为训练样本输入SVM分类器，即可训练得到目标准确的参数化外观模型。由于HOG特征具有几何不变的特性，经过SVM训练后得到的目标外观模型将更加准确。对文中所提出的遥感图像处理算法，实现了仿真程序设计，验证了算法的可行性，完成无人机航空遥感图像动态拼接软件的设计工作。

全书共分10章，第1章对无人机及遥感系统做了简单的介绍，针对遥感技术中的图像

匹配和融合，给予综述性的论述；第2章研究拼接处理前的畸变校正以及去噪增强等图像预处理工作；第3章研究基于图像大边缘提取的图像拼接技术；第4章研究颜色空间和颜色相似性的度量关系，为基于像素颜色的相似性匹配计算提供理论基础；第5章讨论彩色图像拼接；第6章研究目标检测中一些常用的图像特征，详细描述了HOG特征的提取方法；第7章研究支持向量机(SVM)算法的原理，详细描述了基于HOG和SVM的目标外观模型的构建；第8章研究并提出了基于最小生成树模型的目标模型，详细描述了跟踪模型的构建过程，并将这种跟踪算法应用到多目标跟踪视频中；第9章研究一个在线的、结构化的SVM(Structured SVM)框架中学习外观模型参数和目标间的相互关系参数，实时更新所有目标的外观模型和这些目标间的结构约束，使跟踪器能够及时适应目标和环境的变化；第10章总结全文并展望下一步的研究内容。

本书第1章至第5章由程远航编写，第6章至第10章由武刚编写。全书由程远航统稿。

作者的能力水平和学识有限，本书内容如有不足之处，恳请读者批评指正。反馈邮箱：wkservice@vip.163.com。

作者

2016年3月

目录

第1章
绪论



本章对无人机及其遥感系统做了简单的介绍，针对遥感技术中的图像匹配和融合，给予综述性的论述，使读者能对无人机遥感系统及其图像处理有宏观的认识和了解。

1.1 无人机

无人机(Unmanned Aerial Vehicle，UAV)是一个用于描述最新一代无人驾驶飞机的术语。无人机用途广泛，成本低，效费比好；无人员伤亡风险；生存能力强，机动性能好，使用方便，在现代战争中有极其重要的作用，在民用领域有广阔的前景。无人机分为侦察机和靶机。侦察机用于完成战场侦察和监视、定位校射、毁伤评估、电子战等；也可民用，如边境巡逻、核辐射探测、航空摄影、航空探矿、灾情监视、交通巡逻、治安监控等；靶机可作为火炮、导弹的靶标。

最早的无人机开发是在第一次世界大战后，并在第二次世界大战中将无人靶机用于训练防空炮手。第二次世界大战之后将多余或者退役的飞机改装成特殊研究用途飞机或者靶机，近代无人机开始投入使用。近30年来，随着微电子、通信、材料及推进系统等技术的迅猛发展，世界各国开始重视无人机的研究与发展，无人机得到了长足的进步。在越南战争期间，美国曾经使用大量的无人机对高价值或者防御严密的目标进行侦察，以减少人员伤亡的风险。1982年，以色列航空工业公司(IAI)首创以无人机担任其他角色的军事任务。在黎巴嫩战争期间，侦察用无人机曾经在以色列陆军和空军的服役中担任重要角色，完成侦察、情报收集、跟踪和通信等任务。由于无人机成本低廉，灵活机动性强，使用中不存在人员伤亡的风险，因此得到了美国军方的高度重视。最初的无人机以侦察机为主，在1991年的沙漠风暴作战中，美军曾经发射小型无人机作为诱饵，专门用于欺骗敌方的雷达系统。目前，一些无人机已经装备了武器(例如RQ—1捕食者装备、AGM—114地狱火空对地导弹)，装备有武器的无人机被称为无人战斗机(UCAV)，可执行更多的战斗飞行任务。美国军队曾购买或自制先锋无人机在对伊拉克的第二次和第三次海湾战争中完成作战任务。对于非军用无人机的商业兴趣，导致一些企业致力于设计和销售自主飞行器。

目前，世界上已有50多个国家研制出50多种无人机，无人机型号超过300余种，主要用于国防安全和军事侦察等目的，如美国洛克希德·马丁和波音公司联合制造的

Darkstar，美国通用原子能公司的Pradator，以色列飞机工业公司的Heron，以及意大利阿莱尼亚公司的Mirach等。最为著名的是美国研制的“全球鹰”无人机，它是现有无人机中最大的一种，主要用于执行高空、远程和长航时连续监视侦察任务，能近实时地提供高分辨率地面图像。

MQ—8B“火力侦察兵”垂直起降无人机、美国RQ—1“掠夺者”无人机、美国RQ—4“全球鹰”无人机、美制捕食者B无人机，分别如图1-1、图1-2、图1-3、图1-4所示。



图1-1　MQ—8B“火力侦察兵”垂直起降无人机



图1-2　美国RQ—1“掠夺者”无人机



图1-3　美国RQ—4“全球鹰”无人机



图1-4　美制捕食者B无人机

我国的无人机工业起步虽然较晚，但发展迅速。目前，多家科研院所和公司研制了各种类型的无人机，如南京航空航天大学研制的“翔鸟”无人直升机，北京航空航天大学研制的“长虹”机载发射、高速度多用途无人机，中国测绘科学研究院研制的UAVRS—Ⅱ型民用遥感无人机，北京维思韦尔航电有限责任公司推出的“隼鹰—2”“蜂鸟—1”和“鲨鱼—2”等型号无人机以及贵航集团研制生产的大型无人机。国内无人机的研究发展在总体设计、飞行控制、组合导航、中继数据链路通信系统、传感器技术、图像传输、生产制造和部队使用等诸多技术领域积累了一定的经验，具备了一定的技术基础。特别是近几年，一批新型无人机装备相继研制成功并交付使用，无人机装备体系结构也有了较大的改善，但从总体水平上来说，我国无人机装备技术同发达国家相比仍有较大的差距。

1.2　无人机航空遥感系统

由于无人机具有机动灵活、快速反应等特点，以无人机为飞行载体的航空遥感系统，作为一种新型高分辨率对地观测系统，已成为对地观测系统的重要组成部分。无人机航空遥感系统具有可低空飞行、图像分辨率高、续航时间长、图像实时传输、适合高危地区作业、成本低、机动灵活等优点，适用于低空高分辨率的遥感数据的实时获取。在区域性、工程性、灾害性和军事性的遥感监测中发挥着大型遥感系统难以替代的作用，在某些特殊

的应用场合尤为重要。随着成像传感器技术的发展，无人机航空遥感系统所获取的图像数据包括灰度、彩色、视频图像等多种形式。为了满足实际应用的需要，实现图像实时监测，图像监控系统应具备可以自动、高速地完成图像的获取、压缩、传输、处理、显示以及存储等功能，以保证地面接收系统能够迅速、准确地获得遥感图像数据，用于观察、决策以及相关后续处理。在高危地区探测、边防监控、地面目标攻击等应用中，无人机遥感图像处理的精准性、可视性与实时性是无人机能否得以有效利用的决定性因素。

作为一种特殊的飞行载体，无人机的飞行特性复杂、飞行速度较快、成像条件多变，因此具有更为复杂的图像运动特性。无人机多在某一局部地区完成监测任务，在特殊目的下需要对地面目标进行实时跟踪飞行，需要低空、盘旋飞行，飞行姿态多变，飞机与地面景物之间的相对运动显著，成像载荷获得的图像产生较大的畸变和直线、旋转运动模糊。受特殊天气和光照、相机离焦等因素的影响，获取的遥感图像会出现对比度低、模糊甚至遮挡现象。图像质量的降低，不仅使地面人员无法清晰观察地面情况，也不利于图像的后续处理与分析，如目标跟踪和定位等。随着无人机飞行速度、遥感图像的分辨率、数据采样频率和通信频带宽度的不断提高，海量遥感图像数据的自动、高速、高质量的实时处理已成为急需解决的关键问题。针对以上问题进行研究，解决相关的关键技术问题，建立高性能的航空遥感图像处理系统，将为无人机遥感监测系统的广泛应用提供前提与基础，具有重要的学术价值和紧迫的实际意义。

由于无人机特殊的飞行特性和对图像监控系统的特殊功能、性能要求，无人机遥感图像监控系统与一般的图像处理系统相比有所不同，主要表现在以下几个方面：①图像数据采样周期短、分辨率高、数据量大；②对于监测目的遥感图像处理有更高的实时性能要求；③遥感图像传输和处理过程中要求数据损失小，处理精度高。因此，在无人机遥感图像监控系统的构建和实现上都存在很多需要进一步研究和解决的问题，对图像的实时处理和分析这一研究领域带来了一项艰巨挑战。近年来，一些以小型无人机或无人飞艇等为搭载平台的轻小型航空遥感监测系统虽然已经建立，但由于成像条件、处理方法的限制，在成像质量和处理速度方面仍然无法实现高水平和高效能的实时遥感图像监控作业。

针对实时图像处理的客观迫切要求，20世纪70年代以来，国际上逐渐将数字图像处理技术的研究重点放在实时图像处理技术上，这是因为针对动态图像只有在实时处理时才有实际应用价值。20世纪80年代初期，美国陆军夜视电子实验室、海军电子系统司令部、海空研究所、海空研究中心、美国国家宇航局以及一些大型电子公司和高等院校都致力于

开展实时图像处理技术的理论和应用研究。20世纪80年代中期以来，美国已在该领域取得重大进展，其三军的微光和红外观察系统、制导武器、航空航天等都在不同程度上应用了实时图像处理技术。此外，德国、以色列等在图像实时配准处理技术方面也取得了较大的进展，推出了一些相关产品。美国的科研机构在无人机航空遥感图像处理领域起步较早，进行了大量的高分辨率航空遥感图像的科学研究与开发利用工作，并已取得较多的应用成果，如利用无人机遥感高分辨率图像采样能力应对紧急安全事件、农业监测、实时森林火情监测、军事侦察等。

无人机遥感图像实时处理技术，同样得到了我国科研人员的普遍关注，国内多家科研院所、高校都进行了相关的研究工作。2001年3月，中国科学院空间科学与应用研究中心在实验室环境下利用计算机和天线卫星模拟了从图像生成到压缩、加密、GPS复合、发送、接收等实时传输过程，取得了良好的仿真实验效果。2002年6月，中国测绘科学研究院完成无人机遥感监测系统UAVRS—II的研制，实现了机载遥感设备、遥感设备控制系统、地面监控系统的集成，并开发了图像后处理软件。2004年7月，中国国防科技大学通过对遥感图像预处理算法的可并行性分析，提出了对适用于大数据量遥感图像实时预处理的分布并行处理系统的结构设计方案。同年12月，浙江大学地球科学系GIS重点实验室设计并实现了有自主版权的大数据量遥感图像数字处理系统MAMS的算法模块，并提出了优化算法。青岛天骄公司科研人员在袁业立院士的带领下，成功研制了我国首个50公斤级“TJ—1型无人机遥感快速监测系统”，它可在一定程度上满足对突发性事件应急动态监测和小范围大比例尺制图的需要。北京大学与贵航集团共同研制、中科院遥感所协助参与的多用途无人机遥感系统，完成遥感图像系统的集成设计、空中遥感控制子系统以及地面数据接收分发子系统的研制工作，集成遥感成像、数据处理、接口与通信等多方面技术，初步解决了高分辨率数据的获取、存储、传输、显示等基本问题，并于2005年8月在安顺黄果树机场首飞实验成功，标志着我国工业型无人机的遥感技术进入实用阶段。对于无人机遥感图像实时处理的研究工作，我国的研究工作虽然已取得初步的成果，但多以实验室研究或试验系统为主，产品化的成果极少，与发达国家相比仍然存在一定的差距，仍有很多工作需要进一步完善和提高。

1.3 遥感图像拼接技术

无人机在执行遥感监测任务时，在飞行过程中将获取的图像数据和飞行状态参数实时传输给地面，由地面专家对获取的地面图像进行观察和分析。通常情况下，由于成像设备性能所限，无法获得大面积的区域图像，因此需要将获取的序列遥感图像或视频图像实现实时拼接并显示，以保证地面专家全面掌握地面目标情况以做出正确的相应控制决策。给定某个场景的一组互相有重叠的局部图像，生成包含这组局部图像的一幅完整的宽视场图像，称为图像拼接。图像拼接在全景图像的构成、地球卫星照片合成、3D立体模型的构造、遥感图像处理、医学图像分析和计算机视觉等方面都有很重要的应用。一般来说，把要拼接的图像分别称为参考图像和待拼接图像，往往通过寻找待拼接图像与参考图像之间的联系来确定拼接过程中可能会出现的参数。图像拼接是遥感图像应用的重要研究内容之一，特别是对于较大范围的卫星遥感和区域性航空遥感而言，目标区域往往不能由一幅图像所覆盖，因此，需要由多幅图像进行拼接，形成较大范围的遥感图像。图像拼接需要在相邻两幅图像的重叠区域中搜索并匹配相同的图像内容，从而确定两幅图像之间的相对位置关系，实现相邻两幅图像对接处理。从实现方法上划分，图像拼接可分为手动拼接和自动拼接两类。手动拼接容易使拼接后的图像在接缝处出现错位或整体图像的各区域颜色不一致的现象，无法有效地保证图像的质量和拼接工作效率。利用计算机实现图像的自动拼接，可提高拼接的质量和效率，已成为图像拼接处理中的主要研究内容。图像拼接过程中需要进行大量的数值计算，如何实现快速、精确的图像拼接，是当前图像拼接研究领域的一个关键问题。

图像拼接通常可以分成两个步骤：配准(Registration)和融合(Blending)。配准的目的是根据几何运动模型，将图像注册到同一个坐标系中。融合则是将配准后的图像合成一张大的拼接图像。

配准可以分成两个步骤：局部配准和全局配准。局部配准是求解相邻两幅图像间的运动关系，通常是求解二维平面运动关系：对应矩阵。全局配准是处理多幅图像拼接到一起的误差累积以实现多幅图像的精确拼接。融合需要将源图像的像素结合起来生成拼接平面上的像素，实现相邻图像间自然的过渡。

1.3.1 图像匹配技术

1. 基于灰度相关的匹配

基于灰度相关的匹配是在待定点为中心的窗口内，以图像的灰度分布为匹配的内容，故称之为灰度匹配，又因为匹配是以划分的图像块为单位进行的，所以又称为块匹配(Block Matching)。这类算法的性能主要取决于相似性度量准则及搜索策略的选择。相似性度量是指用什么度量来确定待匹配特征之间的相似性，它通常定义为某种代价函数或者距离函数的形式。经典的相似性度量包括灰度差的绝对值平方和(SSD)、灰度差的绝对值和(SAD)和相关函数等。近年，人们又提出了Hausdorff距离(HD)、互信息作为匹配度量。传统的Hausdorff距离对噪声非常敏感，分数Hausdorff距离能处理当目标存在遮挡和出格点的情况，但计算费时，用HD进行图像匹配存在的最大问题是匹配的鲁棒性较差。为此，研究者对经典的Hausdorff距离做了一系列改进以提高匹配的鲁棒性。有学者[15]提出了一种鲁棒型Hausdorff距离，除了利用边缘点的位置信息外，还考虑了其他类型的信息。当图像降质或出现出格点或遮挡等现象时，此距离具有较好的鲁棒性；基于互信息的方法因其对于光照的改变不敏感而在医学图像匹配等领域得到广泛的应用，但也存在计算量大的问题。由于基于灰度匹配算法的计算量大，为了提高计算速度，人们提出了很多搜索方法，其中著名的有：三步法(Tss，Three-step search)[16]、四步法(Fss，Four-step search)[16]、梯度法、二维对数法、交叉法和菱形法等。这些算法通过设计不同的搜索模板和搜索策略，在计算复杂度上比完全搜索减小了很多，但在搜索的准确性上比不上全搜索。有学者[17]对梯度法进行了改进，基本思想是首先对当前块在参考帧中的起始搜索点进行预测，然后再用SGDS(Square Gradient Descent Search)模板进行估计。与传统的搜索方法相比，该方法可减少搜索点数并提高信噪比。匹配窗口大小的选择是基于灰度相关的匹配方法必须考虑的问题，大窗口在遮挡或景物深度不连续的情况下会出现误匹配，小窗口不能覆盖足够的强度变化，因此自适应调整匹配窗口的大小可达到较好的效果。通过比较参考图像和输入图像在各个位置的相关系数，取相关值最大的点就是最佳匹配位置。基于图像灰度信息的匹配方法比较直观，容易实现，并且具有以下优点。

(1) 不需要额外计算抽取图像的特征，可以避免由于特征抽取所引入的参数估计误差。

(2) 可以很好地处理具有丰富纹理特征的区域，而基于特征的匹配算法对具有较强纹理的图像处理效果较差。

(3) 它所利用的信息量是最大的，区分不同对象的能力强、精度高，特别是在十分复杂的图像环境和要求高精度的应用领域仍能够有效地工作。然而，在待检测图像中以模板位置为中心，相似性度量值有时会形成比较平缓的相关曲线，此时要找到准确的匹配位置难度很大。而且，图像的灰度值通常会受到光照条件的影响，非线性的不均匀光照会造成最终结果的较大偏差。同时，图像的比例变化、旋转、遮挡等都会影响匹配结果。此外，基于灰度相关的匹配计算量大，很难达到实时性要求。

2. 基于特征的匹配

为了克服基于灰度相关匹配方法的缺点，人们提出了基于特征的匹配方法。该类方法首先从待配准的图像中提取特征，用相似性度量和一些约束条件确定几何变换，最后将该变换作用于待匹配图像。匹配中常用的特征有边缘、轮廓、直线、点、颜色、纹理等，下面就特征点、线特征和矩特征的匹配进行介绍。

(1) 基于特征点的匹配。人们早先是使用图像的边缘特征进行匹配，但是边缘的计算比较费时，后来又提出用信息量高、数量相对少的特征点代替边缘来计算两幅图像之间的相似性。特征点是图像灰度有较大变化的一类局部点，包含角点、拐点以及T交叉点等。定位精度、重复率和信息含量是衡量特征点检测算子的标准。Moravec[19]在1977年就提出利用图像灰度自相关函数的兴趣点检测算子。Harris等人对Moravec算子进行了改进，提出Harris角点检测算子。Harris算子对于图像存在旋转、照明变化和透视变形时比较稳定，因此得到广泛应用。有学者[20]使用Harris检测器提取特征点，通过变归一化相关，沿极线寻找一幅图像中兴趣点的对应点，接着利用第三幅图像来得到更准确的对应；Jane[21]提出了基于小波变换的分层图像匹配算法，在分解后的每一层图像中提取特征点进行匹配，用并行策略进一步提高速度。最近，Lowe[22]提出的尺度不变的SXRT算子，它在图像发生平移、旋转、缩放及一些光照条件变化的情况下，仍能够保持不变，具有很好的鲁棒性。

(2) 基于线特征的匹配。线特征是图像中明显的线段特征，包括一般的线段、物体的轮廓、道路和海岸线等。线的对应一般表示为线的端点或中点的对应。现有的图像边缘提取方法可以分为三大类：第一类是局部算子法，这类方法属于经典的边缘提取方法，常借助空域微分算子进行，通过将算子与图像卷积完成，这些算子包括Roberts算子、Kirsch算子、Sobel算子、Prewitt算子、Canny算子和LOG算子等；第二类是以能量最小化为准则的全局提取方法，它的特征是运用严格的数学方法对此问题进行分析，给出代价函数作为最优提取依据，从全局最优的观点提取边缘，如最小代价函数法[54][55]、动态规划法[56][57]、神

经网络分析法[58]等；第三类是基于小波变换[59][60][61]、数学形态学[62]等近年发展起来的新理论的图像边缘提取方法。

(3) 基于矩特征的匹配。矩作为图像的一种形状特征，已经广泛应用于计算机视觉和模式识别等领域。有学者[63]在代数不变量的基础上首先提出几何矩的概念，并使用几何矩的非线性组合，推导出具有尺度、平移和旋转不变性的矩不变量。几何矩在信息冗余方面存在欠缺，而且对噪声比较敏感。后来，人们又推导出正交的Zemike矩和Legendre矩。有学者[64]使用矩不变量作为匹配特征，模糊不相似性作为匹配度量，提出了最优匹配理论并加以证明；Huang和Cohentzs提出了一个曲线匹配的算法，使用加权B样条曲线矩，解决仿射变换和遮挡问题。使用矩的匹配方法无须建立点的对应信息，它的缺点是不能检测图像的局部特征，需要对图像进行分割，而且只适用于发生刚体变换的图像。由于图像的特征比像素点在数量上要少很多，因此特征匹配可减少匹配的计算量。同时，特征的匹配度量值对位置的变化比较敏感，使得匹配的精确度较高。通过特征的提取可减少噪声的影响，对灰度变化、图像形变以及遮挡等都有较好的适应能力，所以基于特征的匹配在实际中的应用越来越广泛。需要强调的是，特征匹配和特征提取密切相关，一个稳定、准确的特征提取方法可以提高匹配的精确度，而特征提取引入的计算量也会直接影响匹配算法的效率。

3. 基于模型的匹配

模板匹配方法在计算机视觉和模式识别等领域的应用非常广泛，它可以分为刚体形状匹配和变形模板匹配两大类。在刚体形状匹配中，原型模板通过平移、旋转和尺度化等简单变换达到与目标图像相匹配，但是它不能处理目标形状存在较大变形的情况。为此，人们提出使用变形模板匹配方法。Jain[65]将变形模型分为自由式变形模型和参数式变形模型。Kass提出的Snake模型、主动轮廓模型[66]是比较典型的自由式变形模板模型。Snake是能量极小化的样条，模板不包含目标的全局结构，它只受局部连续性和平滑约束。给定初始轮廓后，它在三种能量的驱动下改变轮廓的位置和形状：约束轮廓形状的内部能量、将轮廓吸引到显著特征的图像力、引导它的行为的外力。当由上述三项构成的能量函数最小化时得到最终的图像特征。由于不受全局结构的限制，所以Snake模型能表示任意的形状，但是该模型对于模板的初始位置和噪声比较敏感，对于凹边缘的收敛性差，而且容易陷入局部最小值。参数化的变形模板使用了几何形状的一些先验信息，模板用少量的参数来表示，典型的例子包括Grenander模型、活动形状模型和Blake的活动轮廓。有两种方法

可以用来表示参数化变形模板：①用参数化曲线的集合表示，模板的几何形状通过改变参数值而改变；②参数映射下的原型模板图像，它通过指定一个原型模板来描述一类目标的形状，该类目标的各种形状都是通过对原型进行参数映射而得到的，不同的参数值可以得到不同的形状。在上述两种表示方法中，变形模板通过图像力动态调整参数达到与图像特征之间的交互作用。与主动轮廓模型相似，参数式变形模板用内部能量和外部能量的加权和构成的目标函数确定一个变形模板和给定图像中目标的匹配程度。使用主动轮廓模型对可见光图像和SAR图像进行匹配。该方法首先提取可见光图像的轮廓作为Snake的初始模板，对该初始模板进行变形从而提取SAR图像的轮廓，之后在提取出来的两幅图像的轮廓上选择控制点计算变换参数。

4. 频域匹配

此方法利用空域到频域变换的良好性质：空域函数平移、旋转和变尺度在频域都有其对应性。对于图像的平移，计算两幅图像功率谱的逆傅立叶变换可得一个脉冲函数，该函数仅在平移量处不为零。对于旋转，可使用极坐标方式表示，用相同方法计算图像间的旋转角度，假如图像运动是平移和旋转的混合运动，可分两步运算：先计算旋转后的平移，经过频域变换，两组数据在空间上的相关运算可以变为在频域上频谱的复数乘法运算。该方法对小平移量和旋转及变尺度图像整合非常适合，并有专用的硬件和快速算法，因此计算速度快，同时，能克服相关性噪音和依赖频率噪音，可适用多传感器和光源变化采集的图像。频域匹配对噪声有较高的容忍度，检测结果与照度无关，可处理图像之间的旋转和尺度变化。另外，由于图像在频域中能获得在空域中难以获得的特征，如傅立叶变换的相位特性在匹配中能得到非常好的输出信噪比和输出最大峰值，比空域有更好的精度和可靠性。除了傅立叶变换外，人们还可以选择更可靠、更符合人眼视觉生理特征的Gabor变换[67]以及小波变换[35]进行图像匹配。在图像匹配中应该结合实际情况选择合适的方法进行匹配，除了上面提到的分类，也会将几种方法综合起来得到更好的结果。在实现的过程中，除了原理，还需要考虑诸如效率、精度等一些其他因素，尤其是搜索策略。搜索策略是在保证不影响匹配结果的同时，用合适的搜索方法在搜索空间中找出平移、旋转等变换参数的最优估计，使得图像之间经过变换后的相似性最大，而花费的时间最少。搜索策略有穷举搜索、分层搜索、模拟退火算法、Powell方向加速法、动态规划法、遗传算法和神经网络等。遗传算法采用非遍历寻优搜索策略，可以保证寻优搜索的结果具有全局最优性，计算量较之遍历式搜索要少很多；神经网络具有分布式存储和并行处理方式、自组织和自学习

的功能以及很强的容错性和鲁棒性，因此这两种方法在图像匹配中得到较为广泛的使用。在成像过程中，由于噪声及遮挡等原因，导致一幅图像中的匹配基元在另一幅图像中有多个候选基元得不到对应，这些都是初级视觉中的“不适定问题”。通常，通过几何约束来解决这种问题，常用的约束有唯一性约束、连续性约束、相容性约束和顺序一致性约束。同时，人们还采用最小二乘法和投票算法等后处理来有效地消除假配点和误配点。

1.3.2 图像融合技术

图像融合的方法有很多，大致上可以分为以下几类：基于颜色空间变换的融合方法，典型的有IHS变换的融合方法；基于多尺度变换的融合方法，典型的多尺度变换有小波变换、拉普拉斯变换等；基于加权平均的融合方法；逻辑滤波方法、数学形态法、模拟退火法等。应该注意的是这几类方法并不是严格区分的，有的方法是几类方法的结合。比如基于多尺度变换的融合方法，在多尺度变换后，可以在变换域上做加权平均进行融合。我们下面将简单介绍这几类融合方法。

1. 基于颜色空间变换的融合方法

IHS变换是这类方法中最常用的颜色空间变换。这种方法多用在高分辨率的SPOT图与分辨率较低但具有丰富颜色细节的TM图之间的融合。它的做法是将TM图从RGB的颜色空间变换到IHS(Intensity Hue Saturation)，称为HS Value、HSI或HIS空间，然后对TM图进行高分辨率重采样，将得到的H和S分量放到SPOT图的相关像素上，而原SPOT图的颜色分量则取为I分量，然后变换回RGB空间，完成融合。这里可能会有一些变化，TM图的I分量与SPOT图可以采取某种方式(如加权平均)进行融合，而不是替代，这种方式在待融合的图像分辨率相同或相近的情况下用得比较多。还有一种更简单的方法，就是在待融合图像的各个颜色分量中选取重要的颜色分量作为结果图像RGB中的某一分量。我们称这种方法为多通道彩色融合法。

2. 加权平均的融合方法

几乎所有的融合方法在最终融合时都可能采用加权平均的融合方法，加权平均可以在空域和变换域上进行。加权平均的融合方法的关键是加权系数的选择，可分为权系数全局一致和自适应的求取权系数方法。PCA融合方法也是一种加权平均方法，它根据多幅图像的协方差矩阵计算其加权系数，权系数对所有像素取为一致。目前，我们更多地采用自适

应加权平均的融合方法。加权平均法的优点是简单直观，适合实时处理。但简单的叠加会使合成图像的信噪比降低；当融合图像的灰度差异很大时，就会出现明显的拼接痕迹，不利于人眼识别和后续的目标识别过程。

3. 基于多尺度变换的融合方法

这类方法是近几年图像融合方法研究的热点，有学者[69]对这类方法进行了归纳总结，给出这类方法的基本框架，以两幅输入图像为例，G.Piella提出的图像融合框架的框图，如图1-5所示。y_A、y_B为输入图像的多分辨率分析系数，虚线框内为融合规则，可分为4部分：活性因子、匹配因子、系数择取和系数合成。活性因子和匹配因子用以提取多分辨率分析系数的特征信息，这些特征信息可用于系数择取，从而得到相应的择取因子，然后利用择取因子通过系数合成得到融合图像的多分辨率分析系数，最后通过多分辨率分析逆运算得到最后的融合图像。

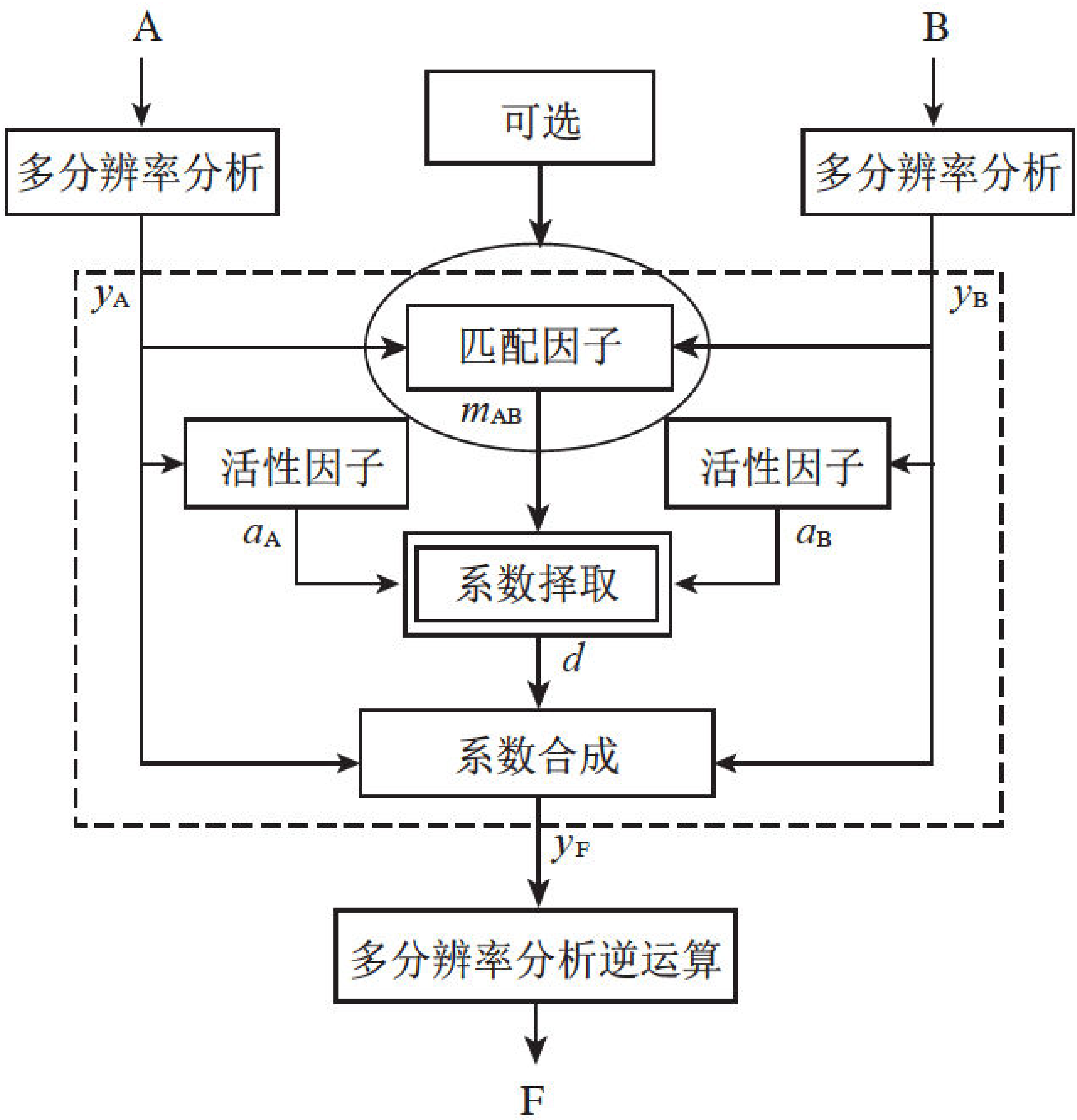


图1-5　图像融合示意图

多尺度融合方法主要有两个步骤：多尺度分解和融合规则。多尺度分解有很多选择，如小波分解、拉普拉斯金字塔分解、梯度金字塔分解等。在完成分解后得到图像的一种多分辨率表示，可以采用金字塔结构来组织，之后的融合在这个结构上进行。不同算法的

融合规则虽然五花八门，但其目的都是突出特征，既要保留图像的边缘信息，也要保持图像原有的光滑性。为了刻画图像的特征和边缘信息，引入活性因子度量(Activity Level Measure)；而为了保证融合图像有一定的光滑性，我们需要考察待融合图像相同位置的相关程度，因此这里我们引入匹配因子度量。通过活性因子和匹配因子来共同决定如何融合图像，通常由它们来决定融合的加权系数。

4. 逻辑滤波方法

逻辑滤波方法是将两个像素数据通过某种逻辑运算合成一个像素的一种直观方法。例如，当两个像素的值都大于某一阈值时，“与”滤波器输出为“1”(为“真”)。图像通过“与”滤波器而获得的特征可认为是图像中十分显著的成分。同样地，“或”逻辑操作能可靠地分割一幅图像。

5. 数学形态法

数学形态法是指通过从基本算子即并、交、差、条件加推演出来的一套数学形态算子如膨胀、腐蚀等，对图像进行处理。若两个集合互相支持，则通过集合从两个特征集中提取出高置信度的“核”特征集；若两个集合互相对抗，则通过一个集从另一个集中提取出高置信度的“潜在”特征集。用条件膨胀和条件腐蚀的形态运算来融合“核”与“潜在”特征集。条件膨胀用来提取“潜在”特征集的连接分量，可用来抑制杂波；条件腐蚀可用来填入在“核”特征集中丢失的分量边界元素。开运算和闭运算的基本作用是对图像进行平滑处理，开运算可以去掉图像中的孤立子域和毛刺，闭运算可以填平一些小洞并将两个邻近的目标连接起来。统计形态的引入为图像融合提供了一种新的思路。将统计的思想与形态滤波相结合估计图像包含的有用信息，噪声抑制效果较好。

6. 模拟退火法

模拟退火法是以张弛为基础的最优化技术。该方法用于图像处理时，相当于把像素值以它们驻留的邻域视为物理系统中的原子或分子的状态，将能量函数分配给物理系统并确定其Gibbs分布。因为Gibbs分布与马尔可夫随机场等价，所以如果图像可以表示为马尔可夫随机场，则能量函数就决定图像模型。能量函数的温度降低，将物理系统退火到全局能量最小状态，则对应于带有噪声的初始图像对真实图像的最大后验估计。模拟退火法用来进行像素层的图像融合，就是要找到能恰当地描述对最终融合图像约束的能量函数。

7. 曝光差异处理方法

该方法分别取不同图像的部分图像数据来合成整幅图像。在连续的图像拍摄过程中，

由于受拍摄条件或曝光时间等因素的影响，不同图像的亮度和颜色分布会存在一定的差异，直接对接两幅图像将会在拼接后的图像中出现明显的拼接缝，使整幅图像出现灰度或颜色不均匀的现象，对拼接图像的视觉效果以及后续的图像处理和分析都将产生影响。为了避免拼接缝的产生，通常采用重叠区域图像数据取平均或采用对重叠区域图像数据渐入渐出的方法来平滑图像拼接缝。[75]这两种方法尽管在一定程度上可以使图像灰度或颜色在拼缝处平滑过渡，局部效果上有所改善，但在图像之间的灰度或颜色差异较大的情况下，图像整体上仍然会存在灰度或颜色分布不均匀的现象。有学者提出了基于小波变换的拼接缝消除方法，以消除图像之间的灰度不均衡现象，这种方法尽管理论严密，但计算量较大，实现困难，很难达到理想的平滑效果。

与灰度遥感图像相比，彩色遥感图像具有更为丰富的可视化信息，因此得到越来越广泛的应用。彩色图像中像素颜色通常需要由三维颜色空间向量来表示，因此与灰度图像相比，在图像的表达方式和处理方法上存在较大的差异，使得彩色图像拼接具有更大的难度。以往的彩色图像拼接方法大多首先将彩色图像变换为灰度图像，然后采用灰度图像拼接的方法进行匹配计算。由于在彩色与灰度图像变换过程中，必然会丢失大量有用的颜色信息，因此难以得到理想的拼接处理结果。彩色图像平滑过程中可以通过对图像的拼接缝做各种颜色过渡处理，但由于图像在整体上存在颜色差异，使平滑后的图像很难保证整体上的颜色一致性。人眼对于彩色具有较高的敏感度，彩色图像平滑同样成为图像拼接中的另一个关键技术环节。由于彩色图像的数据量大，如何提高拼接算法的效率是需要特别考虑的问题。

第2章

遥感图像的基础理论及相关知识



无人机遥感系统获取图像时在成像的过程中，由于受到各种因素的影响会产生几何畸变。几何畸变所带来的误差可以分为内部误差和外部误差。内部误差主要是由于传感器自身的性能、技术指标偏离标称数值所造成的。比如，扫描速度不匀、不同波段相同扫描线的成像时间差、扫描线的非线性和非平行性等使得到的图像失真。外部误差是指遥感器的传感器本身处在正常工作的条件下，由传感器以外的各种因素所造成的误差。比如，传感器的外方位(位置、姿态)变化、地球曲率、地球自转等因素引起的变形误差。所以在对图像进行后续处理之前必须进行几何校正。本章主要介绍几何畸变校正的常用方法，在此基础上提出能够加速遥感图像处理的并行技术校正方案。

在获取图像的过程中，受曝光时间等成像条件的影响，会使获取的图像出现偏亮或偏暗的现象，对比度下降，不能很好地反映地物特征。本章提出一种小波变换与Curvelet变换相结合的图像增强方法，对遥感图像实现了对比度增强运算，同时很好地抑制了噪声，保证遥感图像具有更好的清晰度。

2.1 遥感图像几何畸变校正

遥感图像的几何变形是指图像中的像元在图像坐标系中的坐标与其在地图坐标系等参考系统中的坐标之间的差异，消除这种差异的过程称为几何校正。遥感图像的几何变形可分为静态变形和动态变形两大类。静态变形是指在一幅图像的形成过程中传感器相对于地球表面呈静止状态时所具有的各种变形误差。动态变形是指在一幅图像的形成过程中传感器的运动所造成的图像变形。在静态变形中，可分为内部误差和外部误差两类。内部误差是由传感器结构等因素引起的，如摄像机的焦距变动、像主点偏移、镜头畸变等。外部误差是指遥感器本身处在正常的工作条件下，由遥感器以外的各种因素所造成的误差，如传感器的外方位变化、传感介质的不均匀、地球曲率、地形起伏、地球旋转等因素所引起的误差。

无人机在执行航拍任务中，由于飞行运动会使无人机成像传感器的姿态发生变换，导致获取的遥感图像产生几何畸变。对于同一地物区域，不同的飞行姿态所获取的遥感图像

会存在较大的差异，因此需要对获取的图像进行几何校正，以得到基于同一基准投影面上的遥感图像。根据无人机飞行过程中的成像传感器姿态、航拍时间、辅助导航定位数据及其他各种空中传感器所获取的数据进行融合，完成图像的校正。考虑到无人机的实际飞行情况以及对遥感图像的成像质量和实时性要求，几何校正中只对水平方向上的角度偏移所产生的几何畸变进行校正。在每一幅图像拍摄时的飞行状态数据的基础上，利用飞行角度和正北方之间的角度偏移建立校正变换函数。采用邻近插补法，实现几何校正。[11]

2.1.1　坐标系统

空间坐标系，如图2-1所示，具体阐释如下所述。

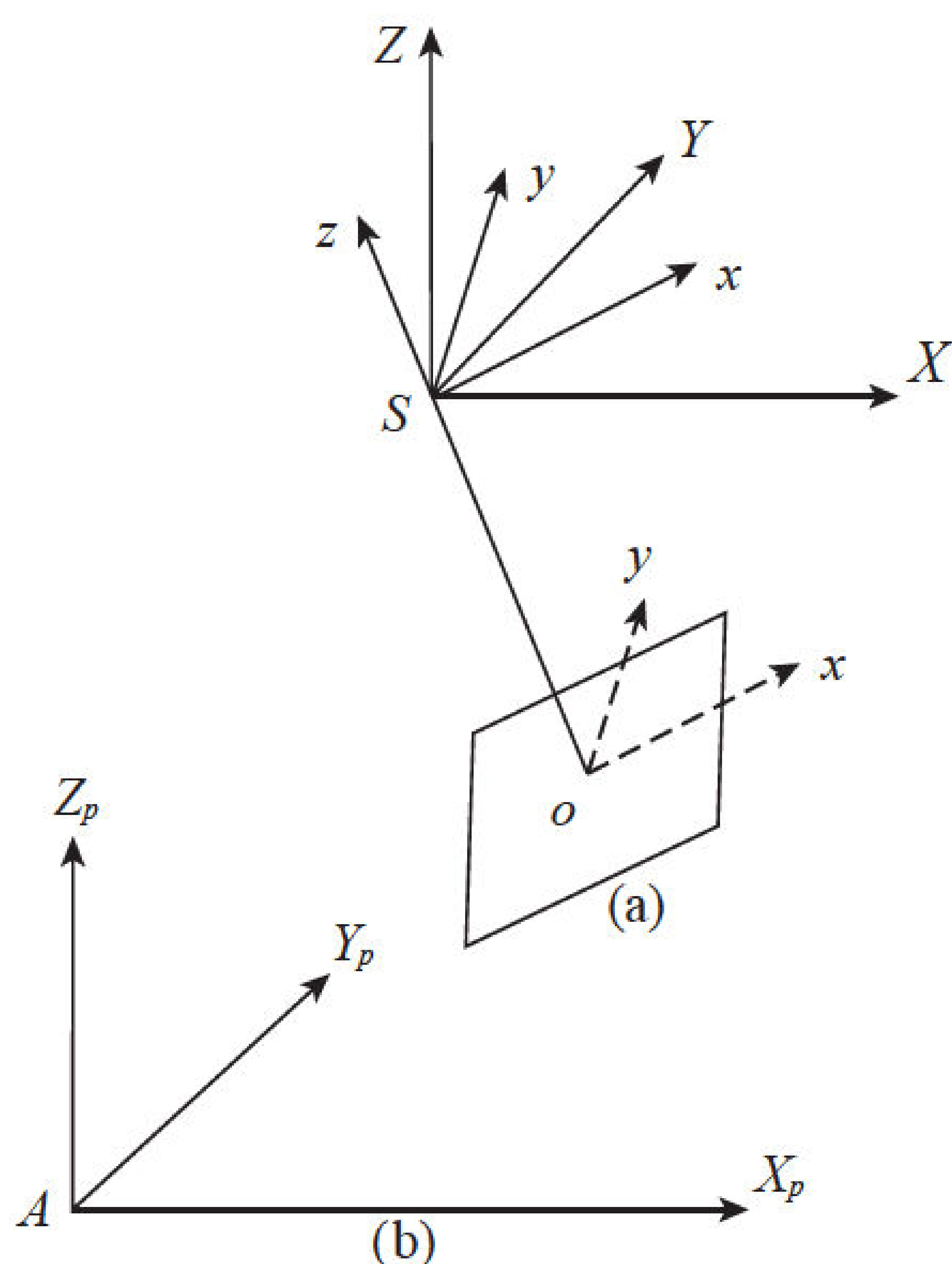


图2-1　空间坐标系

1. 像平面坐标系$o-xy$

像平面坐标系是影像平面内的直角坐标系，用来表示像点在像平面上的位置。若摄影中心为S，摄影方向与影像平面的交点o称为影像的像主点。像平面坐标系的原点就位于像主点。对于航空影像，两对边机械框标的连线为x轴和y轴的坐标系称为框标坐标系，其与航线方向一致的连线为x轴，航线方向为正向，像平面坐标系的方向与框标坐标系的方向相同。

2. 像空间坐标系$S-xyz$

该坐标系是一种过渡坐标系，用来表示像点在像方空间的位置。该坐标系以摄站点(或投影中心)S为坐标原点，摄影机的主光轴So为坐标系的z轴，像空间坐标系的x、y轴平行，正方向如图2-1所示。在这个坐标系中，每一个像点的z坐标都等于So的长，但符号是负的。

3. 像空间辅助坐标系$S-XYZ$

该坐标系是一种过渡坐标系，它以投影中心S为坐标原点。在航空摄影中通常以铅垂方向(或设定的某一竖直方向)为Z轴，取航线方向为X轴，这样有利于改正航线方向积累的系统误差。

4. 摄影测量坐标系$A-X_pY_pZ_p$

该坐标系是一种过渡坐标系，用来描述解析摄影测量过程中模型点的坐标。在航空摄影测量中通常以地面上某一点A为坐标原点，而它的坐标轴与像空间辅助坐标系的坐标轴平行。

5. 物空间坐标系$O-X_tY_tZ_t$

该坐标系是摄物体所在的空间直角坐标系。测绘中所用的是地面测量坐标系(大地坐标系)。前面介绍的4种坐标系均是右手直角坐标系，而地面测量坐标系为左手坐标系，它的轴指向正北方向，高程则以我国黄海高程系统为基准。在地球上一个小范围内讨论问题时，把O□$X_tY_tZ_t$视为左手直角坐标系是允许的，但当测区范围较大时，须考虑地球曲率的影响。

2.1.2　影像的内外方位元素

1. 内方位元素

确定摄像机的镜头中心(严格地说，应该是镜头的像方节点)相对于影像中心的位置关系参数，称为影像的内方位元素。内方位元素包括以下三个参数：像主点(主光轴在影像面上的垂足)相对于影像中心的位置x_0、y_0以及镜头中心到影像面的垂距f(也称主距)，如图2-2所示。对于航空影像，x_0、y_0即像主点在框标坐标系中的坐标。内方位元素值一般由摄影机检校确定。

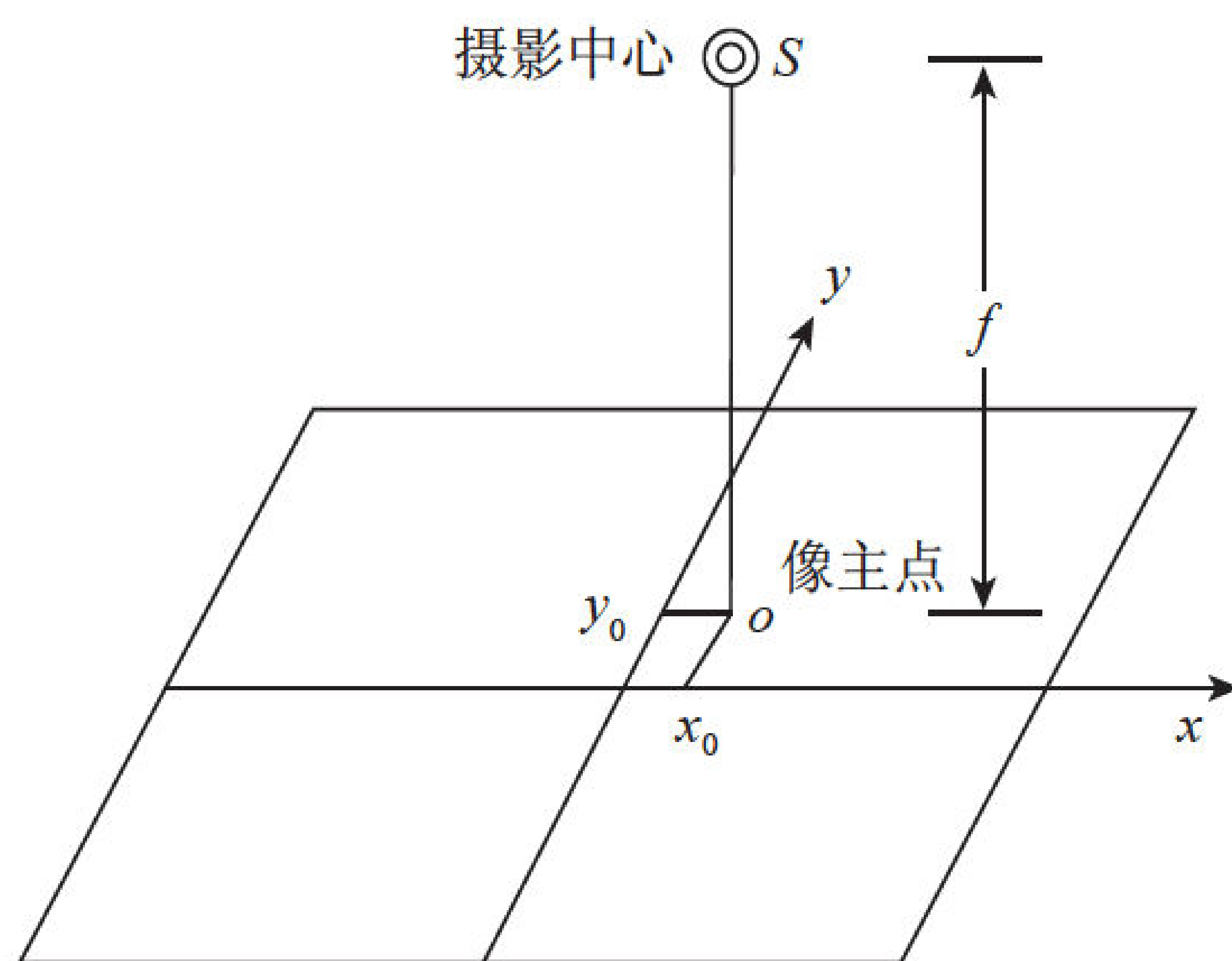


图2-2　影像的内方位元素

2. 外方位元素

确定影像或摄影光束在摄影瞬间的空间位置和姿态的参数，称为影像的外方位元素。一幅影像的外方位元素包括6个参数，其中有3个是线元素，用于描述摄影中心S相对于物方空间坐标系的位置X_S、Y_S、Z_S；另外3个是角元素，用于描述影像面在摄影瞬间的空中姿态。角元素有三种不同的表达形式。①以Y轴为主轴的φ□ω□κ系统(主轴是在旋转过程中空间方位不变的一个固定轴)：以Y轴为主轴旋转φ，然后绕X轴旋转ω，最后绕Z轴旋转κ角。②以X轴为主轴的ω□φ□κ'系统：以X轴为主轴旋转ω'，然后绕Y轴旋转φ'，最后绕Z轴旋转κ'角。③以Z轴为主轴的A□α□κ系统：以Z轴为主轴旋转A角，然后绕Y轴旋转α，最后绕Z轴旋转κ角。

2.1.3　空间直角坐标系的旋转变换

像点空间直角坐标的旋转变换是指像空间坐标与像空间辅助坐标之间的变换。空间直角坐标的变换是正交变换，一个坐标系按某种顺序依次旋转三个角度即可变换为另一个同原点的坐标系。

设像点a在像空间坐标系中的坐标为(x，y，□f)，而在像空间辅助坐标系中的坐标为(X，Y，Z)，两者之间的正交变换关系可以用下式表示

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = R \begin{bmatrix} x \\ y \\ -f \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \begin{bmatrix} x \\ y \\ -f \end{bmatrix} \tag{2-1}$$

或

$$\begin{bmatrix} x \\ y \\ -f \end{bmatrix} = R^T \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{2-2}$$

式中，R为一个3×3阶的正交矩阵，由9个方向余弦组成。以影像外方位角元素φ□ω□κ系统为例，对于上述两个坐标系之间的转换关系可以理解为：像空间坐标系是像空间辅助坐标系(相当于摄影光束的起始位置)依次绕相应的坐标轴旋转φ、ω、κ三个角度以后的位置。此时R可表示为

$$R = R_\varphi R_\omega R_\kappa = \begin{bmatrix} \cos\varphi & 0 & -\sin\varphi \\ 0 & 1 & 0 \\ \sin\varphi & 0 & \cos\varphi \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\omega & -\sin\omega \\ 0 & \sin\omega & \cos\omega \end{bmatrix} \begin{bmatrix} \cos\kappa & -\sin\kappa & 0 \\ \sin\kappa & \cos\kappa & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$$

将R的表达式乘积结果列出后，可得

$$\begin{cases} a_1 = \cos\varphi\cos\kappa - \sin\varphi\sin\omega\sin\kappa \\ a_2 = -\cos\varphi\sin\kappa - \sin\varphi\sin\omega\cos\kappa \\ a_3 = -\sin\varphi\cos\omega \\ b_1 = \cos\omega\sin\kappa \\ b_2 = \cos\omega\cos\kappa \\ b_3 = -\sin\omega \\ c_1 = \sin\varphi\cos\kappa + \cos\varphi\sin\omega\sin\kappa \\ c_2 = -\sin\varphi\sin\kappa + \cos\varphi\sin\omega\cos\kappa \\ c_3 = \cos\varphi\cos\omega \end{cases} \tag{2-3}$$

2.1.4 共线方程

令S为摄影中心，在某一规定的物方空间坐标系中其坐标为(X_S，Y_S，Z_S)，M为任一物方空间点，它的物方空间坐标为(X，Y，Z)。m为M在影像上的构像，相应的像空间坐标和像空间辅助坐标分别为(x，y，□f)和(X_m，Y_m，Z_m)，如图2-3所示。

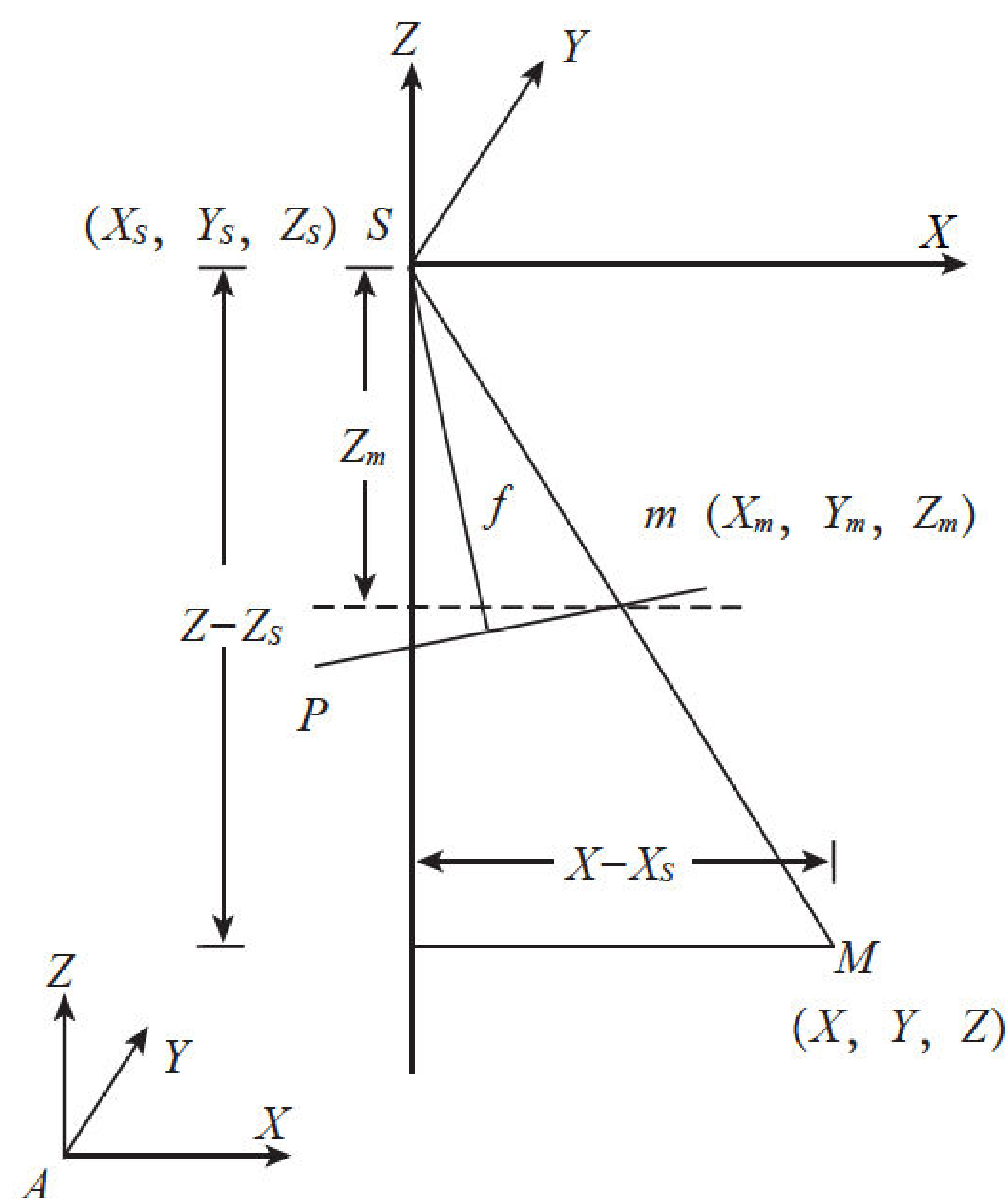


图2-3　共线条件方程

摄影时S、m、M三点位于一条直线上，那么像点的像空间辅助坐标与物方点、物方空间坐标之间有以下关系

$$\frac{X_m}{X-X_s}=\frac{Y_m}{Y-Y_s}=\frac{Z_m}{Z_s-Z}=k \tag{2-4}$$

则

$$X_m=k(X-X_s),\ Y_m=k(Y-Y_s),\ Z_m=k(Z-Z_s)$$

由式(2-1)可知，像空间坐标与像空间辅助坐标有下列关系

$$\begin{bmatrix} x \\ y \\ -f \end{bmatrix}=\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}\begin{bmatrix} X_m \\ Y_m \\ Z_m \end{bmatrix} \tag{a}$$

将式(a)展开为

$$\begin{aligned} \frac{x}{-f}&=\frac{a_1X_m+b_1Y_m+c_1Z_m}{a_3X_m+b_3Y_m+c_3Z_m} \\ \frac{y}{-f}&=\frac{a_2X_m+b_2Y_m+c_2Z_m}{a_3X_m+b_3Y_m+c_3Z_m} \end{aligned} \tag{b}$$

将式(2-4)代入式(b)中，并考虑到像主点的坐标(x_0，y_0)，得

$$
\begin{cases}
x - x_0 = -f\dfrac{a_1(X - X_S) + b_1(Y - Y_S) + c_1(Z - Z_S)}{a_3(X - X_S) + b_3(Y - Y_S) + c_3(Z - Z_S)} \\
y - y_0 = -f\dfrac{a_2(X - X_S) + b_2(Y - Y_S) + c_2(Z - Z_S)}{a_3(X - X_S) + b_3(Y - Y_S) + c_3(Z - Z_S)}
\end{cases}
\tag{2-5}
$$

式(2-5)即为常见的共线条件方程式(简称共线方程)，式中

$(x，y)$为像点的像平面坐标；

$(x_0，y_0，f)$为影像的内方位元素；

$(X_S，Y_S，Z_S)$为摄站点的物方空间坐标；

$(X，Y，Z)$为物方点的物方空间坐标；

$a_i，b_i，c_i(i=1，2，3)$为影像的3个外方位元素组成的9个方向余弦。

由式(2-4)和(2-1)可以推导出共线方程的另一种形式

$$
\begin{bmatrix} X - X_S \\ Y - Y_S \\ Z - Z_S \end{bmatrix} = \frac{1}{k}\begin{bmatrix} X_m \\ Y_m \\ Z_m \end{bmatrix} = \frac{1}{k}R\begin{bmatrix} x \\ y \\ -f \end{bmatrix}
\tag{2-6}
$$

令 $\lambda_m = \dfrac{1}{k}$，并完整地写出旋转矩阵，则有

$$
\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \lambda_m \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}\begin{bmatrix} x \\ y \\ -f \end{bmatrix} + \begin{bmatrix} X_S \\ Y_S \\ Z_S \end{bmatrix}
\tag{2-7}
$$

2.1.5 灰度重采样

常用的像元灰度重采样方法有以下几种。

(1) 双线性插值法。双线性插值法的卷积核是一个三角函数，可以表达为

$$W(x)=1□|x|，0\leqslant|x|\leqslant 1 \tag{2-8}$$

任意像点$p(x，y)$位于4个像元$p_{i,j}$，$p_{i,j+1}$，$p_{i+1,j}$，$p_{i+1,j+1}$之间，则由双线性插值得出其灰度值为

$$
\begin{aligned}
g(x,y) = g_{y,x} &= (1-dx)(1-dy)g_{i,j} + dx(1-dy)g_{i,j+1} \\
&+(1-dx)dy\,g_{i+1,j} + dx\,dy\,g_{i+1,j+1}
\end{aligned}
\tag{2-9}
$$

式中，$dx=x$□$\mathrm{INT}(x)$，$dy=y$□$\mathrm{INT}(y)$，INT为取整部分。

(2) 双三次卷积法。以三次样条函数

$$\begin{cases} \omega_1(x)=1-2x^2+|x|^3 & 0\leqslant|x|<1 \\ \omega_2(x)=4-8|x|+5x^2-|x|^3 & 1\leqslant|x|\leqslant 2 \\ \omega_3(x)=0 & 2<|x| \end{cases} \tag{2-10}$$

为卷积核，需要16个原始像素参与计算。此时

$$g(x,y)=g_{y,x}=\sum_{i=1}^{4}\sum_{j=1}^{4}\omega_{i,j}g_{i,j} \tag{2-11}$$

式中，$\omega_{i,j}=\omega(x_i)\omega(y_i)$

$$\begin{cases} \omega(x_1)=-dx+2dx^2-dx^3 \\ \omega(x_2)=1-2dx^2+dx^3 \\ \omega(x_3)=dx+dx^2+dx^3 \\ \omega(x_4)=-dx^2+dx^3 \\ \omega(y_1)=-dy+2dy^2-dy^3 \\ \omega(y_2)=1-2dy^2+dy^3 \\ \omega(y_3)=dy+dy^2+dy^3 \\ \omega(y_4)=-dy^2-dy^3 \end{cases} \tag{2-12}$$

$$dx=x-\mathrm{INT}(x),\quad dy=y-\mathrm{INT}(y),\quad g_{i,j}=\mathrm{g}(x_j,\ y_i) \tag{2-13}$$

(3) 最邻近像元法。直接取与$P(x,y)$点位置最近像元$N(x_N,y_N)$的灰度值为重采样值，即$g(P)=g(N)$，式中$x_N=\mathrm{INT}(x+0.5)$，$y_N=\mathrm{INT}(y+0.5)$。

上述三种方法中最邻近像元法最简单，计算速度快但精度较差。双三次卷积法采样中误差约为双线性内插的1/3，但计算量大、较费时，因而一般情况下常采用双线性插值法。

2.2　遥感图像并行几何校正算法

随着航天与传感技术的飞速发展，一方面，通过遥感手段获得的信息量大大增加，主要表现在空间分辨率、时间分辨率、光谱分辨率和辐射分辨率的不断提高，单幅遥感图像的数据量已达到数百兆。另一方面，遥感应用不断得到推广，用户不断提出新的需求，很

多领域，如气象预报、灾难监测等，需要对遥感图像进行快速及时、实时的处理。目前对海量数据的快速处理，已成为遥感领域中亟待解决的问题。

采用并行技术是加速遥感图像处理的有效手段，Pbdua等人采用指令集并行技术对遥感图像进行处理，但这种技术与具体硬件的相关性大，通用性差。机群系统性能价格比高，可扩展性好，日益得到并行研究者与用户的青睐。Meisl等基于Cluster(机群)环境实现了SAR信号成像的并行处理，分析了网络带宽对并行性能的影响。Yang在BEOWULF工作站机群系统上进行遥感图像处理，主要考虑了工作站机群环境的搭建，并未提及并行算法的研究，为了加快几何校正的处理过程，针对遥感图像与分布存储并行环境，研究并实现了一种基于局部输出区域计算的并行几何校正算法。

2.2.1 并行几何校正算法

采用并行处理技术是提高算法、执行效率的有效手段，遥感图像的几何校正可采用数据的并行处理方式。[9]下面我们首先介绍直观并行算法及其缺陷，然后针对分布存储的并行环境，提出一种基于局部输出区域计算的并行几何校正算法。

1. 直观并行算法

直观并行算法是大图像并行重采样处理常用的算法[10][11]，其处理方式是对输入图像的存储进行规则(条状或块状)划分，每个处理器保存一块输入子图像，如图2-4(a)所示；同时对目标图像的重采样计算采用简单的规则划分，每个处理器对一块输出子图像进行灰度值重采样，如图2-4(b)所示。

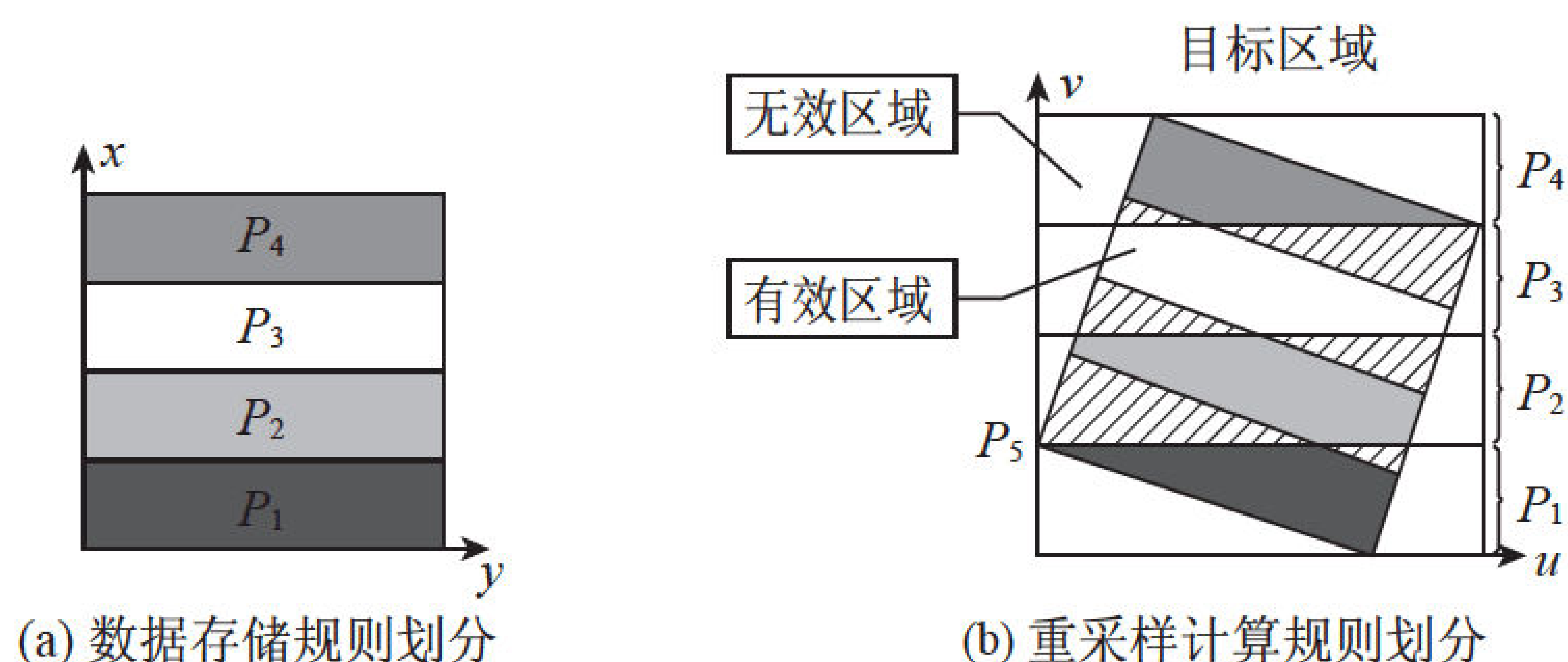


图2-4 直观并行算法数据划分

直观并行算法多用于图像的缩放与微小变形，若在分布存储环境下采用直观并行算法

对卫星遥感图像进行几何校正，则对输出图像块进行重采样计算所需的像元灰度信息可能不在本地，见图2-4(b)中的斜线阴影部分，需要通过处理器之间的交互获得，且这种通信为频繁的小数据传输，开销很大。另外，这种方法对无效区域，即图2-4(b)中的未填充区域的像元需要进行几何位置反算，实际上若知道无效像元的位置，直接将它们的灰度值赋成背景色即可。因此，直观并行算法不适用于分布存储环境下遥感图像的几何校正。

2. 基于局部输出区域计算的并行算法

针对分布存储的并行环境，可提出一种基于局部输出区域计算的并行几何校正算法。它与直观并行算法的不同之处是，每个处理器对输出图像的重采样范围不是规则的图像块，而是一块特定的非规则区域，即各处理器根据正映射关系计算出的本地输入子图像在校正后图像中的范围。考虑到灰度二维插值需要利用邻域像元的信息，为了避免处理器之间的通信，各处理器根据插值模板的大小在本地输入子图像的边界处冗余存储若干行像元，如对于三次卷积重采样，其模板大小为4×4，边界处需要冗余存储两行像元。这样并行重采样计算时各处理器所需的像元灰度值均保存在本地，整个计算过程无须通信，加快了并行处理的速度。负责全局操作的处理器为管理结点，参与并行计算的处理器为计算结点，算法的处理流程见图2-5，其中方框外的操作由管理结点完成，方框内的操作由各计算结点并行执行。

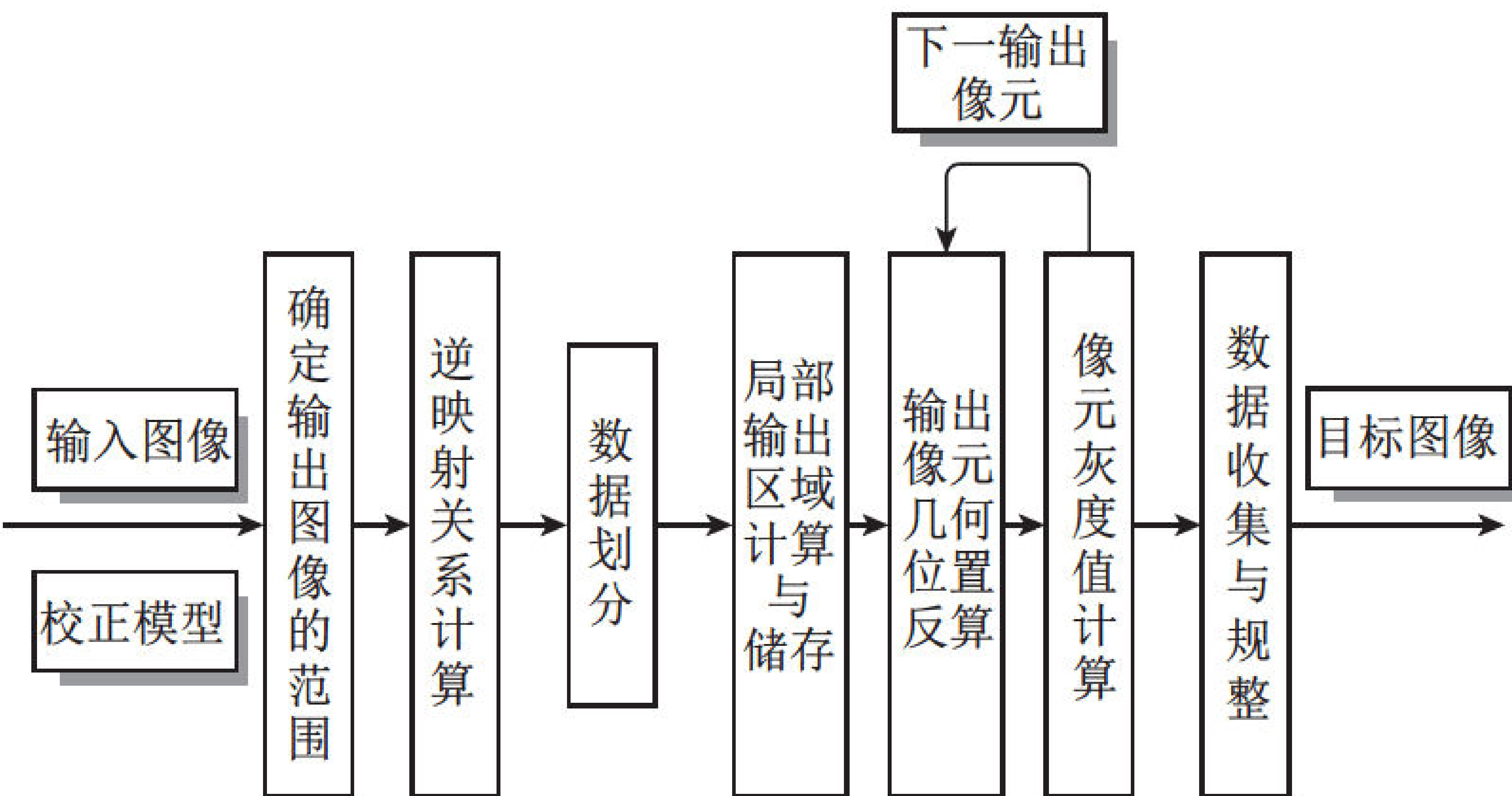


图2-5　并行几何校正算法流程

下面为几何校正算法中各步骤的具体描述。

(1) 确定输出图像的范围与逆映射关系计算。这个步骤与串行算法中的处理方式相同，由管理结点完成，并将求得的输出图像范围及逆映射函数发送到各计算结点。

(2) 数据划分。根据计算结点的数目p对输入图像(畸变图像)进行规则划分，每个计算

结点保存一块输入子图像，并对子图像的边界处进行信息冗余存储。

(3) 局部输出区域计算。计算按本地保存的规则输入子图像对应在目标图像中的范围，确定本结点需要完成的重采样计算，同时采用一种新的存储结构保存该结点的输出图像块信息。这是该并行算法的难点，也是解决数据局部性的关键。

(4) 局部重采样。本地输入子图像对应在目标图像中的范围已知，对于该范围内的每个输出像元，利用逆映射关系求得其在畸变图像中的共轭点，然后利用二维插值确定其灰度值。

(5) 数据收集与规整。各计算结点将本地的结果数据传给管理结点，管理结点按照常规的图像格式，将这些输出图像块进行规整，对于无效区域内的像元点，直接将其灰度值赋成背景色，最后形成一幅完整的校正后图像。

在我们的并行几何校正算法中，局部输出区域的计算与存储保证了算法的数据局部性。如图2-6所示，输入图像被分割成p=4块子图像，在计算结点P_3上，规则输入子图像$abcd$经过正映射函数变换到目标图像的对应区域$a'b'c'd'$，对$a'b'c'$d′范围内包含的全部像元点进行灰度值计算，所需的输入像元信息全部落在本地保存的$abcd$子图像中(包括边界处冗余存储的数据)，从而解决了数据局部性问题。

3. 输出图像块的存储结构

对于无人机遥感图像，一般通过计算畸变图像中4个角点对应的地面坐标就能够确定目标图像的大小。如图2-6中，输入图像的4个角点A、B、C、D对应输出图像中A'、B'、C'、D'4个点，对于所有的有效输出像元，A'、B'、C'、D'4个点中必然存在两个点在v方向上分别取最大值和最小值，这一特点使局部输出区域的计算与存储相对于复杂映射而言较为简单。

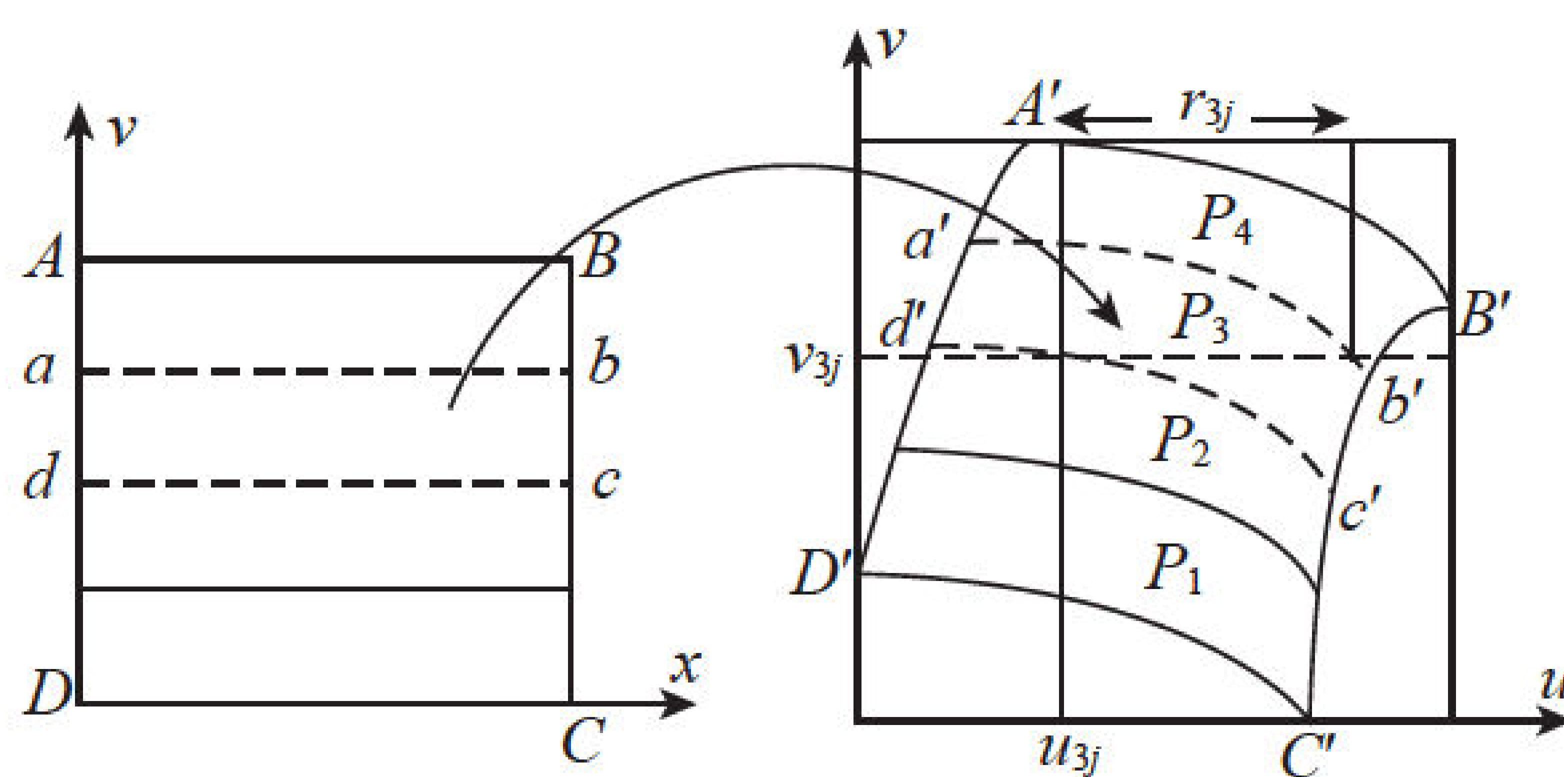
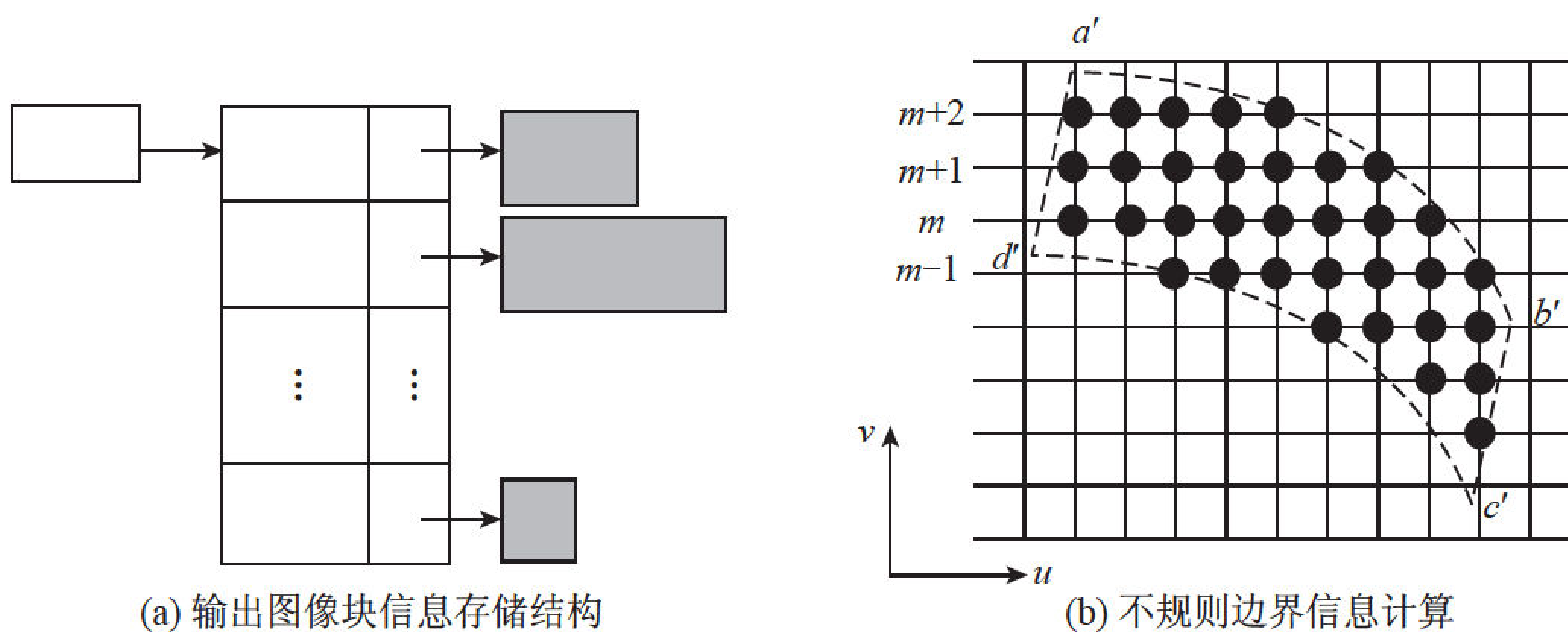


图2-6　数据局部性示意图

首先，我们给出左/右边界段的定义。对于校正后的遥感图像，令其有效区域中v坐标

值最小的点为e，最大的点为f，从点e开始，按顺时针方向绕有效区域的闭合边界一周，则对于从e到f的边界段ef，有效输出像元均在其右部，称为左边界段；从f到e的边界段fe，有效输出像元均在其左部，称为右边界段。同理，对于每个计算结点上的输入子图像，其输出图像块在v方向上也有本地最大值和最小值，相应地有本地左、右边界段。在每条与本地左、右边界段相交的扫描线上，均存在一条像元段，该像元段的起始坐标为扫描线与左边界的交点，像元段的长度为左、右边界段与扫描线形成的两交点之间的距离。图2-6显示在计算结点P_3上，第j条扫描线$v=v_{3j}$与输出图像块的左边界段$d'a'$和右边界段$b'c'$相交，形成的像元段的起始横坐标为u_{3j}，长度为r_{3j}。显然，只需要保存本地输出图像块中每条像元段的起始坐标与长度，就能获得不规则输出图像块的范围，如图2-7(b)所示。为了方便应用，我们将像元段的起始横坐标与长度分别记录为该像元段上起始像元的横坐标及像元的数目。输出图像块的存储结构如图2-7(a)所示，存储的内容包括边界信息与像元灰度值，其中图像块与像元段的结构分别为



(a) 输出图像块信息存储结构　　(b) 不规则边界信息计算

图2-7　输出图像块存储结构及其边界信息计算

```
Struct ImageBlock
{
int   StartVertical;
int   BlockHeight;
struct PixelSeg**PixelSegInfo;
}
Struct PixelSeg
```

```
{
int  StartHorizontal;
int  SegWidth;
int * GrayValue;
}
```

对于每个计算结点$i(1\leqslant i\leqslant p)$，首先记录本地输出图像块的头信息，内容为该图像块在目标图像中跨越的扫描线行数L_i，起始行的横坐标v_{i1}以及包含L_i个元素的指针数组，每个指针指向相应扫描行上的像元段信息。每段像元有一个头信息，包含该像元段中起始像元的横坐标$u_{ij}(1\leqslant j\leqslant L_i)$、像元个数$r_{ij}$以及指向像元灰度值的指针。

4. 局部输出区域计算

令第$i(1\leqslant i\leqslant p)$个计算结点的输入子图像为$abcd$，则其对应的输出图像块$a'b'c'd'$的边界信息$L_i$、$v_{i1}$、$u_{i,j}$、$r_{i,j}$的计算过程如下所述。

(1) 计算边界同名点，确定左、右边界段。根据校正模型，依次计算并保存输入子图像$abcd$上各边界像元点对应在目标图像中的同名点，并称其为边界同名点。令输入子图像的4个角点a、b、c、d对应在目标图像上的位置为a'、b'、c'、d'，点e、f分别为a'、b'、c'、d'中v坐标值最小与最大的点，即$v_e=\min\{va', vb', vc', vd'\}$，$v_f=\max\{va', vb', vc', vd'\}$。在输出图像块中，从$e$点开始，按顺时针方向利用闭合线段依次连接相邻同名点，将形成的闭合边界近似表示输出图像块的理想边界，则边界ef为左边界段，边界fe为右边界段，在图2-7(b)中，$e=c'$，$f=a'$，边界$d'a'$与$b'c'$分别为左、右边界段。

(2) 确定输出图像块信息，$v_{i1}=\lceil v_e\rceil$, $L_i=\lfloor v_f\rfloor-\lceil v_e\rceil+1$，令$v_{i0}=\lceil v_e\rceil-1$，标识量$j$=0。

(3) 求第j+1行扫描线上各边界信息项的值，过程描述如下所述。

① 若$j\geqslant L_i$，边界计算过程结束。

② $j=j+1$，$v_{ij}=v_{i,j-1}$

③若左边界段ef上存在边界同名点s，满足$v_s=v_{ij}$，则$u_{ij}=\lceil u_s\rceil$；否则在ef的边界同名点中找到两个点s、s'，满足$v_s<v_{ij}<v_{s'}$，且无其他同名点S^*，使得$v_s<v_{s^*}<v_{s'}$，则

$$u_{ij}=\left\lceil \frac{v_{ij}-v_{s'}}{v_s-v_{s'}}.[u_s-u_{s'}]+u_{s'}\right\rceil \tag{2-14}$$

④ 若右边界段fe上存在边界同名点t，满足$v_t=v_{ij}$，则$r_{ij}=\lfloor u_t\rfloor-u_{ij}+1$，否则在$fe$上找到两个同名点$t$、$t'$，满足$v_t<v_{ij}<v_{t'}$则

$$r_{ij}=\left\lceil \frac{v_{ij}-v_{t'}}{v_t-v_{t'}}.[u_t-u_{t'}]+u_{t'}\right\rceil-u_{ij}+1 \tag{2-15}$$

(4) 转步(3)。该局部输出区域计算方法的思想是：首先确定L_i及v_{i1}，v_{i0}是为了方便计算引入的初始值，然后依次将v坐标值加1得到v_{ij}，再利用左、右边界段上v值与v_{ij}相邻的两个边界同名点的u、v坐标，用线段拟合输出图像块的理想边界，求出扫描线v=v_{ij}与左、右边界段的交点，从而确定该扫描线上像元段的起始横坐标与该像元段的长度。考虑到有效像元为不规则边界内的像元，因此对u坐标值取整时，像元段左边的u值取上整，右边则取下整。各计算结点利用保存的边界信息即可得到本地输出图像块中各有效像元的位置，然后通过几何位置反算与二维灰度插值获得它们的灰度信息。最后，管理结点对各计算结点的输出图像块结果进行收集规整，形成最终的目标图像。由于目标图像中任意有效像元必须为某个不规则输出图像块的内部像元，因此该算法保证了规整后图像的完整性。

2.2.2　并行几何校正算法的实验结果与分析

我们在机群系统上实现了算法，由于图像的大小是影响算法性能的关键因素，因此在实验中，我们将一幅遥感图像重采样成大小不同的多幅图像，分别进行系统几何校正，分析测试结果。测试平台的配置为：16台微机构成Cluster结构(1个管理结点，15个计算结点)，每个结点的配置为Pentium4，2.2G的CPU，1GB的内存，100Mbps的快速交换式以太网互连，软件环境为Windows2000操作系统，NT2MPICH消息传递并行库。像元的灰度级别为256(8bits)，逆映射关系为完全三次多项式，二维灰度插值利用三次卷积法。采用最小二乘法求逆映射函数时，选取的粗网格在x与y方向上的间隔均为100个像元点。数据由管理结点分发到各计算结点，计算结束后，各输出图像块结果汇集到管理结点进行规整。图2-8给出了实验图像在几何校正前后的缩略图。

由于不断接收数据，因此畸变图像可以源源不断地输入，我们分析在这种流水状态下算法的性能。管理结点负责的全局操作的计算量很少，可完全与数据传输重叠，因此并行几何校正处理的总时间取决于输入与输出图像的传输时间与并行计算的时间。首先，我们分析这两部分时间的理论结果，由于求局部输出区域所需的时间很少，并行计算的时间仅考虑本地重采样时间。假设输入图像有N_1个像元，校正后图像的有效区域有N_2个像元，网络传输速率为V(Mbps)，计算结点的数目为p，则数据传输时间与并行计算时间分别为

$$T_{\text{transfer}} \approx \frac{[N_1 + N_2] \times 8}{10^6 \times V},$$
$$T_{\text{parallel}} \approx T_{\text{local Resanpling}} = \frac{N_2 \times \text{FP}}{p \times \text{FPMA}} \tag{2-16}$$

其中FPMA为各计算结点每秒能够完成的浮点乘加操作数，FP为重采样计算时每个有效输出像元所需的浮点乘加操作数。

当数据达到完全流水，一幅图像进行几何校正所需的时间为$T_{\text{correction}}[p]=\max[T_{\text{transfer}}, T_{\text{parallel}}]$。随着计算结点数的增加，并行计算所需的时间不断减少，而数据的传输时间基本不变(随着冗余数据量的增加会略有增加)，因此存在$p^*(p^* \geqslant 1)$，使得$p<p^*$时，有$T_{\text{parallel}}>T_{\text{transfer}}$；当$p \geqslant p^*$时，$T_{\text{parallel}} \leqslant T_{\text{transfer}}$，为了测试流水状态下算法的性能，我们将同一幅图像反复作为输入图像。表2-1给出了串行几何校正与计算结点数p=8和p=15时并行几何校正的处理时间，表2-2给出了并行处理时的一些详细结果。在我们的实验环境下，对于任意计算结点数$p(1 \leqslant p \leqslant 15)$，全局计算时间均小于并行计算时间与数据传输时间。从表2-2可以看出，当p=8时，$T_{\text{parallel}}>T_{\text{transfer}}$，几何校正的总时间取决于并行计算时间；当$p$=15时，$T_{\text{parallel}}<T_{\text{transfer}}$，总时间取决于数据传输时间。在并行计算中，计算局部输出区域的时间占并行计算时间的百分比随着p的增大而增大，但始终没有超过2%，这说明局部输出区域计算对算法的性能影响不大。

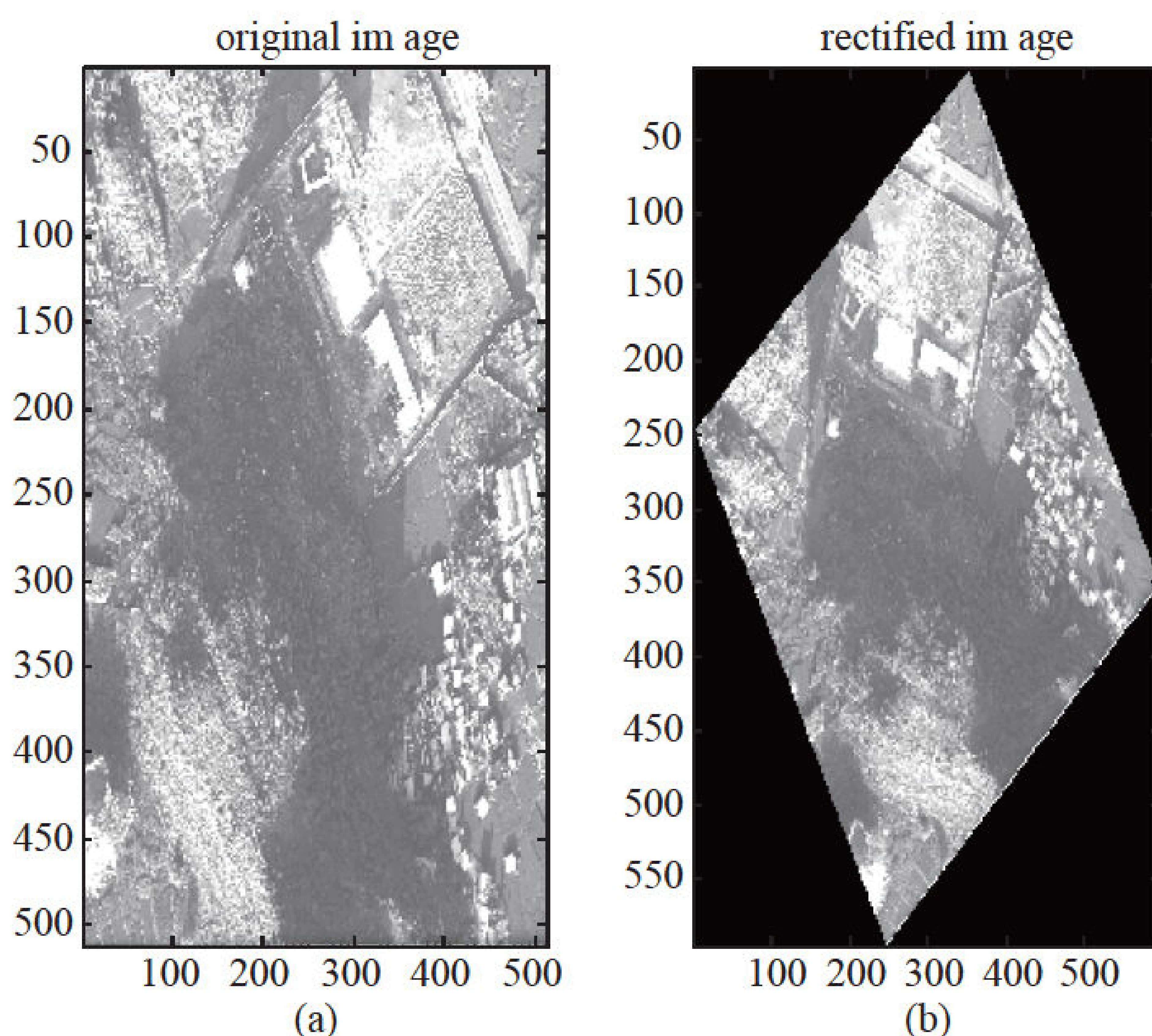


图2-8　遥感图像几何校正结果示意图

表2-1　串行与并行几何校正的实验结果

输入图像大小	串行算法的时间/s	并行计算		并行算法的时间/s			
				数据传输		总时间	
		P=8	P=15	P=8	P=15	P=8	P=15
2002	1.69	0.21	0.11	0.19	0.20	0.21	0.20
6002	3.83	0.82	0.45	0.74	0.76	0.82	0.76
10002	6.77	3.48	3.00	3.52	3.56	7.00	6.56
20002	17.78	10.90	9.17	11.04	10.11	20.07	19.28

表2-2　并行几何校正的详细结果

输入图像大小	全局计算时间/s		局部输出区域计算时间/s	
	最小二乘	规整(P=8)	P=8	P=15
2002	0.002	0.021	0.003	0.002
4002	0.005	0.078	0.011	0.008
10002	0.019	0.781	0.089	0.066
20002	0.038	2.125	0.240	0.180

图2-9给出了采用不同数目的计算结点获得的加速比。从加速比曲线可以看出，对于1000×1000和2000×2000的输入图像，p^*=9；对于两幅大的输入图像，p^*=10。当$p<p^*$时，数据传输时间小于并行计算时间，各计算结点的性能可以得到充分发挥，并行处理的性能由T_{parallel}决定；当$p \geqslant p^*$时，数据传输时间大于并行计算时间，每个计算结点完成一个任务时，需要等待下一幅输入子图像传过来才能再次执行，中间存在等待时间。一旦数据传输时间占主导地位，增加计算结点不能进一步提高并行性能。图2-9显示算法的加速比可能大于计算结点的数目，其原因是串行处理时需要对目标图像中的无效像元进行几何位置反算，而并行算法无须对这些像元进行处理，这在一定程度上减少了运算量。

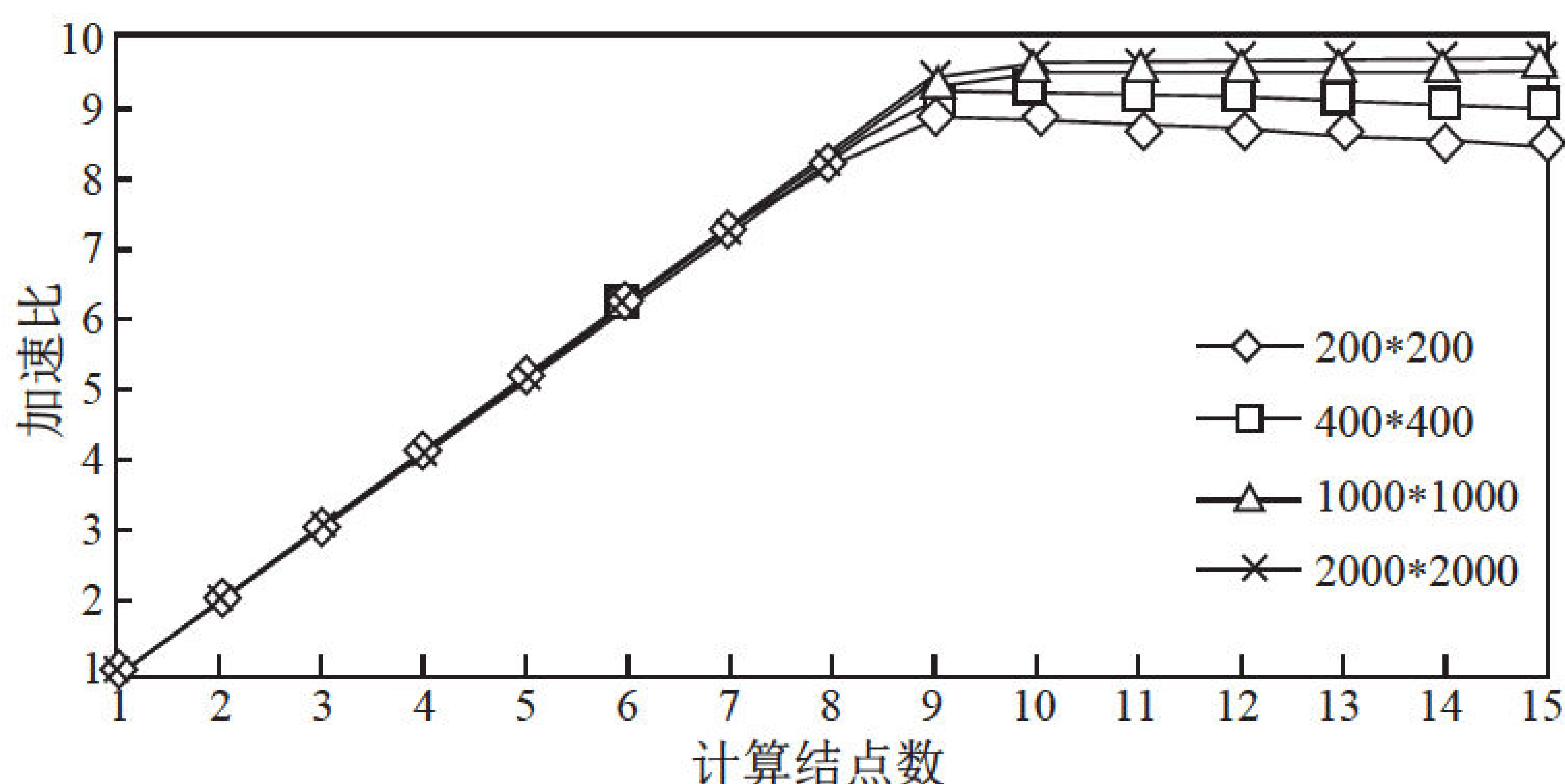


图2-9　不同计算结点数获得的并行加速比

通过上述分析可知，当$p<p^*$时，并行几何校正算法具有很好的并行性能，当$p \geqslant p^*$时，算法性能会随着计算结点数的增加而略有降低。显然，p^*越大，可以利用更多的计算结点获得更好的性能，算法的可扩展性也越好。对于给定的几何校正任务，增大p^*的有效方法是降低数据传输时间，即提高网络带宽。

2.3 遥感图像对比度增强

无人机遥感系统获取图像的过程中，受到曝光时间等成像条件的影响，会使获取的图像出现偏亮或偏暗的现象，对比度下降，不能很好地反映地物特征，因此需要对遥感图像实现对比度增强运算，保证遥感图像具有更好的清晰度。常用的图像对比度增强方法主要包括直方图修正法和直接灰度变换法。[30]下文首先介绍这两类方法。

2.3.1 直方图修正法

数字图像的直方图是作为图像每一个灰度级的统计概率分布，提供了图像灰度分布的概貌。直方图修正法是利用调整图像灰度分布来提高图像的对比度，使图像更加清晰。直方图修正法又可分为直方图均衡化和直方图规定化两种方法。直方图均衡化是指将原图的直方图通过变换函数修整为平坦的直方图，然后按均衡直方图修正原图像，使修正后的图像灰度分布更加均匀，图像更加清晰。直方图规定化也称为直方图匹配，该方法可使处理后的图像具有特殊规定的直方图。在直方图均衡化方法的基础上，又产生了局部区域修正方法，是指针对图像中的特殊局部区域进行直方图修正。[36]下文将具体介绍直方图均衡化方法。

令变量r和s分别代表图像增强前后的像素灰度级，相应灰度级分布的概率密度分别为$P_r(r)$、$P_s(s)$。[13]理想情况下，图像经直方图修正后，将使所有灰度级上的像素分布概率密度为1。设原像素的灰度值r被归一化到区间[0，1]，且概率密度为$P_r(r)$；经过变换后的像素灰度为s，概率密度为$P_s(s)$，变换函数为$T(r)$，则有以下等式

$$s=T(r), \quad 0 \leqslant r \leqslant 1 \tag{2-17}$$

$$P_s(s)\mathrm{d}s=P_r(r)\mathrm{d}r \tag{2-18}$$

变换函数为$T(r)$必须同时满足以下条件：

(1) $T(r)$是单值单调递增函数，以确保其反变换的存在，保证灰度的单值映射，使变换后图像的亮度范围不会被颠倒；

(2) $0 \leqslant T(r) \leqslant 1$，保证图像经过变换后不会缩小灰度动态范围。

如果令

$$s=T(r)=\int_0^r p_r(w)\mathrm{d}w \tag{2-19}$$

则有

$$\frac{\mathrm{d}s}{\mathrm{d}r}=\frac{\mathrm{d}\boldsymbol{T}(\boldsymbol{r})}{\mathrm{d}r}=\frac{\mathrm{d}}{\mathrm{d}r}\left[\int_0^r p_r(w)\mathrm{d}w\right]=p_r(r) \tag{2-20}$$

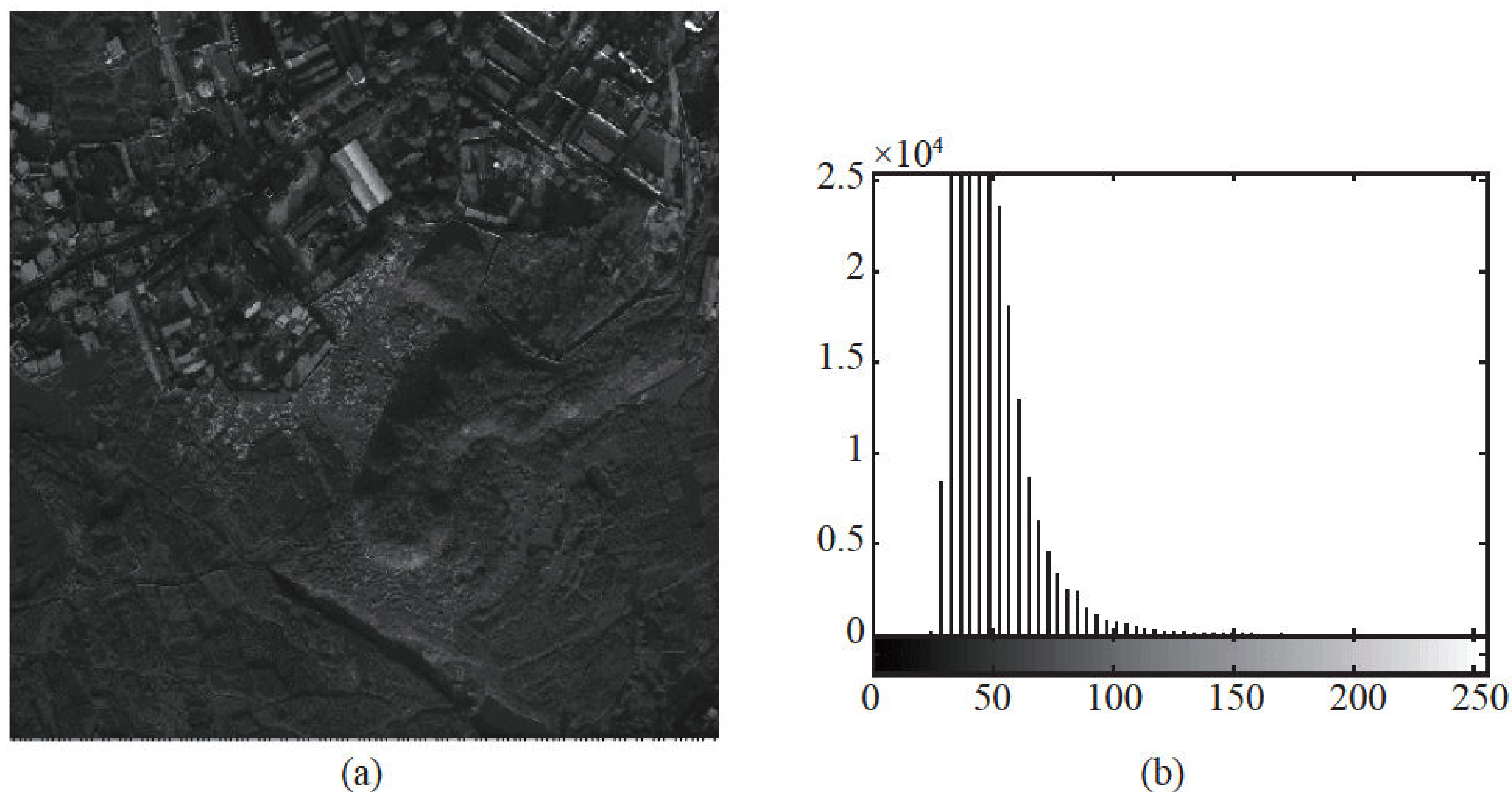
将式(2-20)代入式(2-18)可得

$$p_s(s)=p_r(r)\left|\frac{\mathrm{d}r}{\mathrm{d}s}\right|=p_r(r)\left|\frac{1}{p_r(r)}\right|=1 \quad 0\leqslant s\leqslant 1 \tag{2-21}$$

由此可知，当变换函数取式(2-21)时，则变换后的概率密度函数为常量1。以上是连续函数变换的公式，用于数字图像处理时，如果图像灰度有L阶，则变为

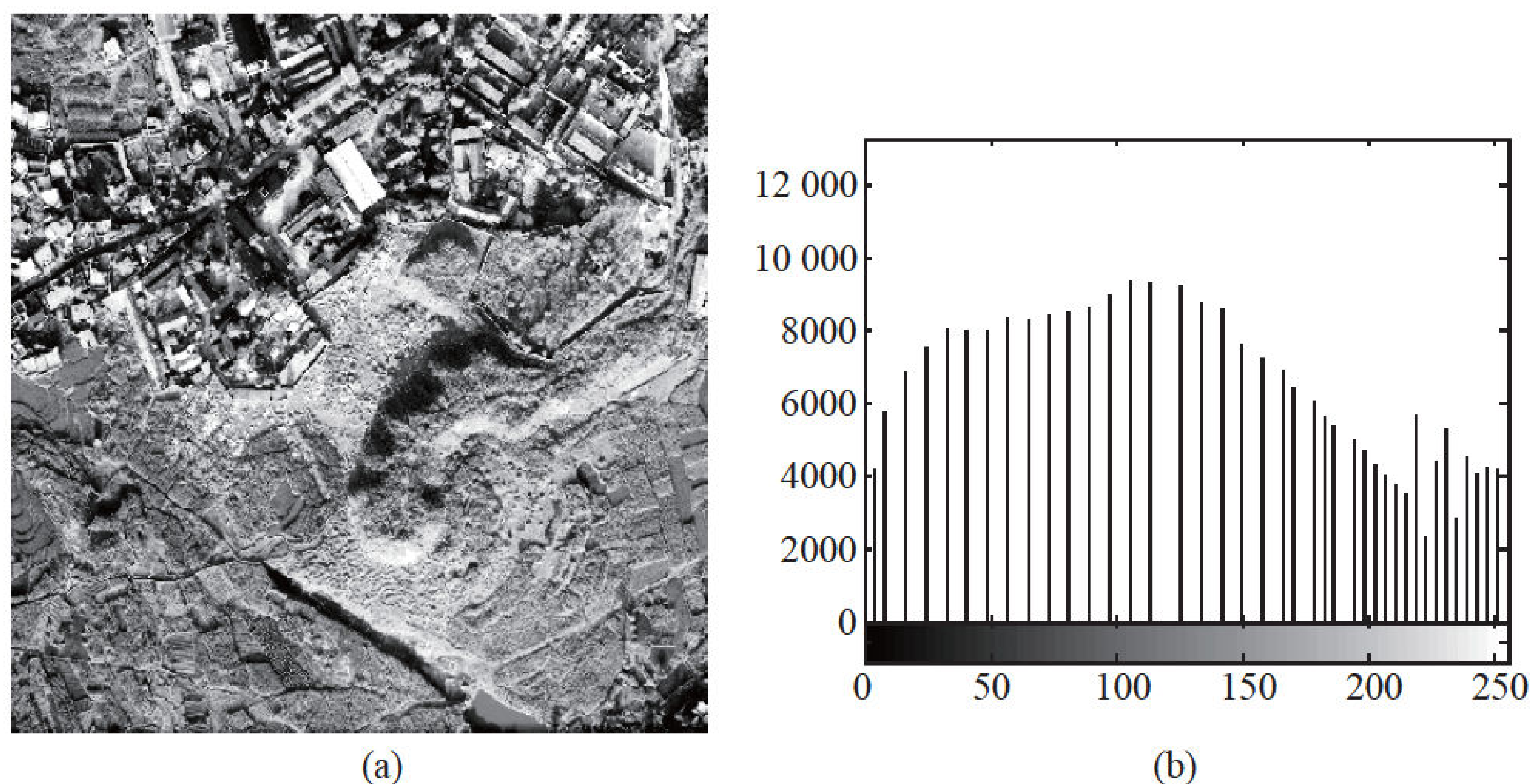
$$s_k=T(r_k)=\sum_{j=0}^{k}p_r(r_j)=\sum_{j=0}^{k}\frac{n_j}{n} \quad k=0,1,2,\cdots,L-1 \tag{2-22}$$

式中，k代表图像的灰度级数，n代表总像素数，n_j代表j灰度层上像素的个数，$p_r(r_j)$代表j灰度层上的概率密度，$T(r_k)$代表k灰度层上像素的变换函数，$s_k(0\leqslant s_k\leqslant 1)$为最终的变化结果。原始图像及直方图，如图2-10所示。均衡化后的图像及直方图，如图2-11所示。



(a)　　(b)

图2-10　原始图像及直方图



(a)　　(b)

图2-11　均衡化后的图像及直方图

2.3.2　直接灰度变换法

直接灰度变换法是比较简单、常用的增强图像对比度的方法，也称为对比度变换、点运算或点处理。常用的直接灰度变换法包括灰度级线性拉伸和灰度级非线性拉伸。[36]这些方法使整幅图像或图像中某一区域里的各像素点的灰度级，都按某一规律进行变换，由此可改变图像的对比度。可用下式表示

$$g(x, y)=T[f(x, y)] \tag{2-23}$$

式中，$g(x, y)$、$f(x, y)$分别表示对应于像素在(x, y)的理想灰度级、实际灰度级，$T[\square]$是一个预定的变换函数，不同的函数关系可以得到不同的变换结果。

1. 灰度级线性拉伸

假定原图像的灰度级范围是$[a, b]$，变换后图像的灰度级范围扩大到$[m, n]$，那么实现这一目标的线性变换公式为

$$g(x, y)=\frac{n-m}{b-a}[f(x, y)-a]+m \tag{2-24}$$

式中，$g(x, y)$、$f(x, y)$为变换前后(x, y)点的灰度值，一般m、n可分别取0、255。

在图像处理过程中，可以采用灰度变换的另一种形式，即分段线性变换。

$$g(x,y)=\begin{cases}\dfrac{c}{a}\times f(x,y) & 0\leqslant f(x,y)<a\\ \dfrac{d-c}{b-a}[f(x,y)-a]+c & a\leqslant f(x,y)\leqslant b\\ \dfrac{M_g-d}{M_f-b}\times[f(x,y)-b]+d & b<f(x,y)\leqslant M_f\end{cases} \tag{2-25}$$

式中，M_f、M_g分别为原图像和变换后图像的最大灰度级。如图2-12所示，通过调节折线拐点的位置及各折线的斜率，即控制参数a，b，c，d的取值，可实现对任一灰度区间的扩展和压缩。分段线性的处理结果如图2-13所示。经过分段线性变换后，可使关心的图像细节的灰度范围得以扩展，增强其对比度；同时使不关心的图像细节处的灰度范围得以压缩，降低其对比度。灰度变换前后，整幅图像总的灰度范围是不变的。

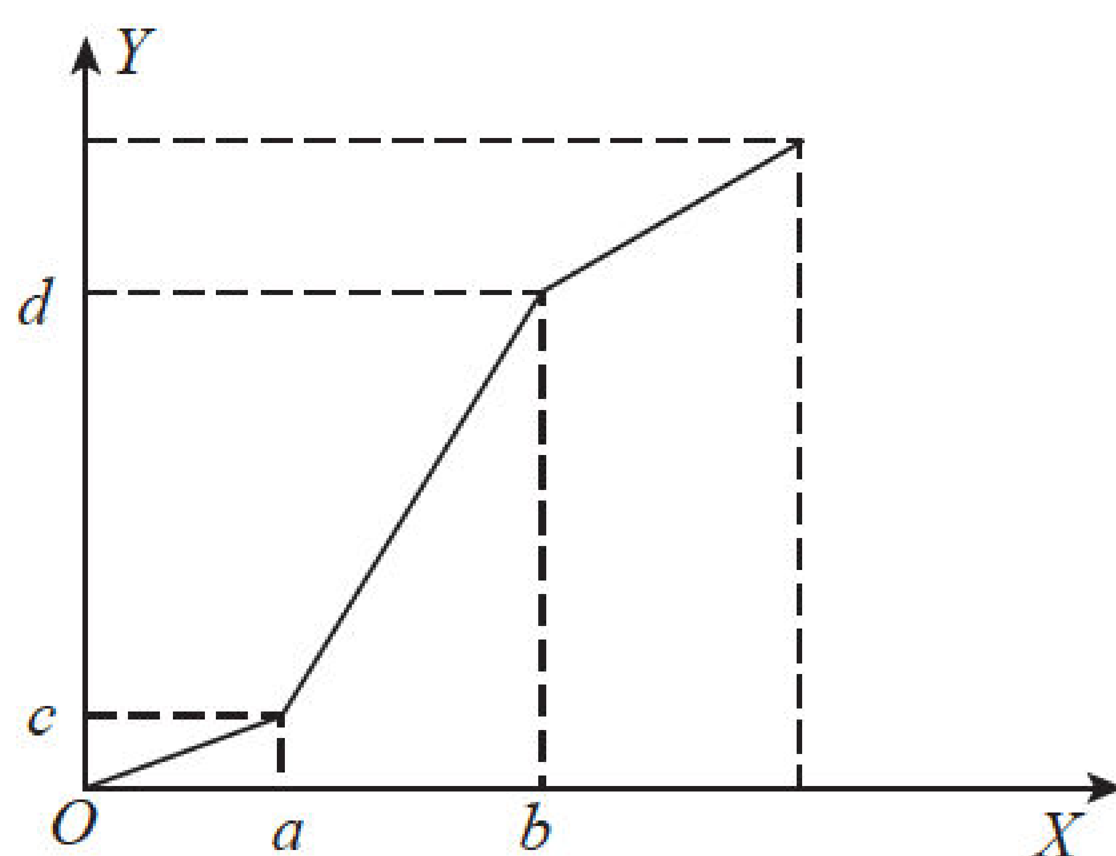


图2-12　分段线性变换



(a) 原始图像



(b) 线性拉伸后的图像

图2-13　灰度线性拉伸增强结果

2. 灰度级非线性拉伸

非线性变换是指希望按照要求的非线性关系对原图像进行变换。有时，需要在某一灰

度级范围内扩展较多，在该灰度范围两端扩展较少，使增强后的图像视觉效果更好，这时就可以使用灰度级非线性变换。

如图2-14所示，可以利用一些数学函数进行非线性变换，如对数、平方、指数等。对数变换的一般表达式为

$$s=c\log(1+r) \tag{2-26}$$

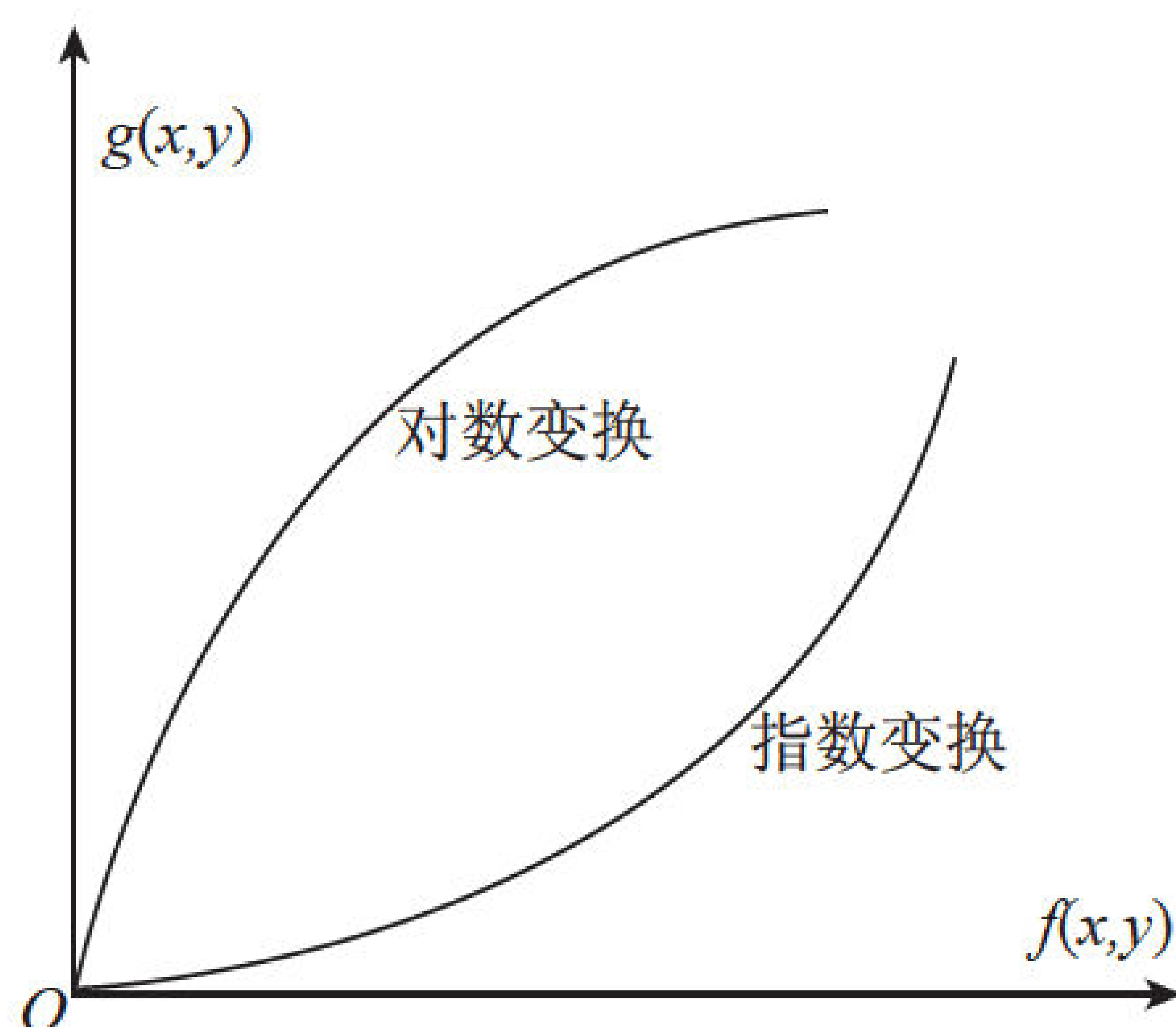


图2-14　非线性变换

其中c是一个常数，并假设$r\geqslant 0$，可以利用变换来扩展被压缩的高值图像中的暗像素。

相对的是反对数变换的调整值。幂次变换的基本形式为

$$s=cr^{\gamma} \tag{2-27}$$

其中c和γ为正常数。有时考虑到偏移量(即当输入为0时的可测输出)，式(2-27)也可写成$s=c(r+\varepsilon)^{\gamma}$。处理结果如图2-15和图2-16所示。



(a) 原始图像



(b) 对数变换后的图像(c=35)

图2-15　对数变换增强结果



(a)　原始图像



(b)　幂次变换后的图像(c=2，γ=1.002)

图2-16　幂次变换增强结果

2.3.3　Wavelet与Curvelet变换相结合的图像增强方法

近年来，小波分析(Wavelet)理论受到众多学科的共同关注。小波变换是传统傅里叶变换的继承和发展。由于小波的多分辨率分析具有良好的空间域和频率域局部化特性，对高频采用逐渐精细的时域或空域步长，可以聚焦分析对象的任意细节，因此特别适合于图像信号这一类非平稳信源的处理，已成为一种信号/图像处理的新手段。Curvelet变换是继小波变换之后，更适合图像处理特点的一种多尺度变换，它能同时获得对图像平滑区域和边缘部分的稀疏表达，且具有很强的方向性，已有初步结果显示其在图像处理中的发展潜力。Velde提出了一种基于小波变换的弱边缘增强算法。首先采用二进制小波对图像进行分解，再对小波系数进行修正，以此达到边缘增强的目的。但基于小波变换的对比度增强方法对于各向异性的特征检测不具有自适应性，所以在对非直线的边缘增强时就不能达到理想的效果，而Curvelet变换由于采用了具有高度方向敏感性的基元，在对各种不同形状的边缘进行增强的时候能起到很好的效果。根据两种增强方法各自的优点，提出基于小波和Curvelet变换相结合的增强算法(CEA)对图像进行增强处理。

1. 小波增强方法

与傅里叶变换相比较，小波变换拥有空域(时域)和频域“变焦距”分析能力。对二维

函数$f(m, n)$进行离散小波变换，能够得到一个低通子带S_j和三个具有方向选择性的高通子带图像$H_{i,j}$，对低通带S_j继续分解可以得到$f(m, n)$的多级小波分解。[9]这一分解过程的结果可以表示为

$$W\{f(m, n)\}=\{(H_{i,j}=W_{i,j}[f(m, n)], S_J\} \quad 1\leqslant i\leqslant 3, 1\leqslant j\leqslant J \tag{2-28}$$

式中，$H_{i,j}$是指尺度j和方向i上的细节信息；S_J是指在最高尺度j上对图像的概貌描述，代表了图像的低频信息。

图像的高频内容集中在低尺度的高通带$H_{i,j}$中，而图像中噪声部分主要集中于低尺度的$H_{i,j}$中。为了克服普通增强方法中常有的对噪声过分增强的缺点，应该对不同尺度的分解系数进行不同程度的增强。

2. Curvelet增强方法

离散Curvelet变换过程如图2-17所示，从图中我们可以看出Curvelet变换可分为以下4个步骤。[10]

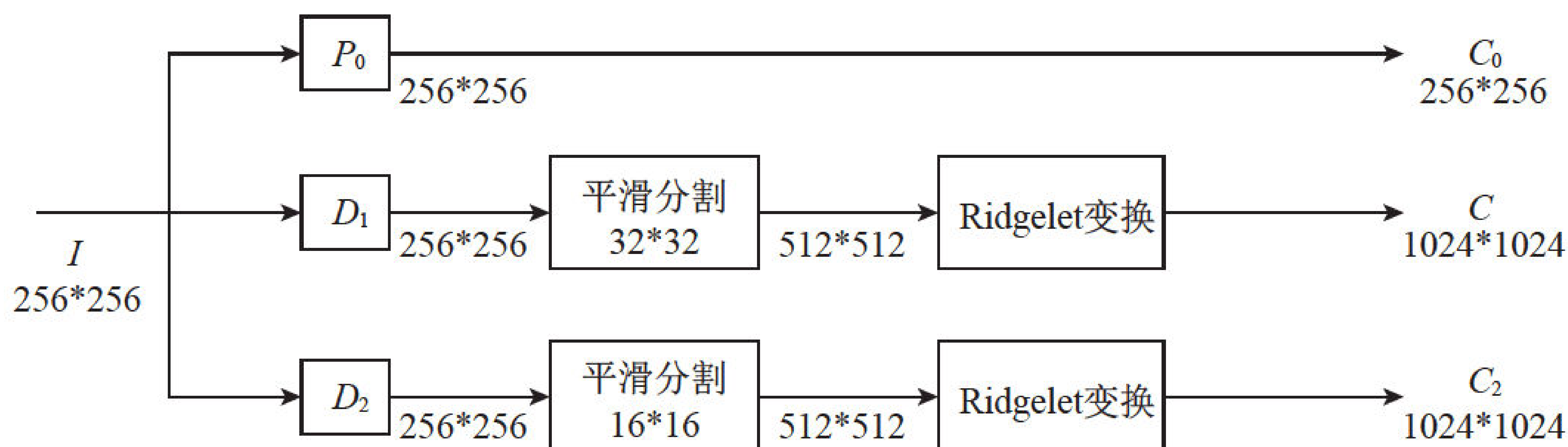


图2-17 Curvelet变换过程

1) 子带分解

在Curvelet变换的分解过程中，首先要进行子带分解，采用滤波器p_0，$D_s(s\geqslant 0)$将I分解为

$$I\,|\!\rightarrow (p_0 I, D_1 I, D_2 I\ldots) \tag{2-29}$$

2) 平滑分割

为了实现平滑分割，定义平滑窗口$w_Q(x_1, x_2)$，该窗口位于方形区域

$$Q=[\frac{k_1}{2^s}, \frac{k_1+1}{2^s})\times[\frac{k_2}{2^s}, \frac{k_2+1}{2^s}) \tag{2-30}$$

将窗口与函数相乘，就得到位于Q的方形区域。分割过程应用于I的各个子带

$$D_s I \big| \to (w_Q D_s I)，\ Q \in Q_s \tag{2-31}$$

3) 重正规化

对每个方形区域Q进行定义

$$(T_Q I)(x_1，\ x_2)=2^s I(2^s x_1 - k_1，\ 2^s x_2 - k_2) \tag{2-32}$$

对上述平滑分块进行重正规化后得到

$$g_Q=(T_Q)^{-1}(w_Q D_s I)，\ Q \in Q_s \tag{2-33}$$

4) Ridgelet分解

最后，进行基函数为ρ_λ的正交Ridgelet分解，分解过程如图2-18所示。首先对图像进行Radon变换，然后再执行一维小波变换。由于Curvelet变换采用了具有高度方向敏感性的基元，所以在对各种不同形状的边缘进行增强的时候能起到很好的效果。在Curvelet增强算法中，定义一个函数y_c对Curvelet变换的系数进行修正。

$$\begin{aligned}
&y_c(x，\sigma)=1，\text{if}(x< c\sigma)\\
&y_c(x，\sigma)=\frac{x-c\sigma}{c\sigma}\left(\frac{m}{c\sigma}\right)^p+\frac{2c\sigma-x}{c\sigma}，\ \text{if}(x<2c\sigma)\\
&y_c(x，\sigma)=\left(\frac{m}{x}\right)^p，\ \text{if}(2c\sigma\leqslant x<m)\\
&y_c(x，\sigma)=\left(\frac{m}{x}\right)^s，\ \text{if}(x\geqslant m)
\end{aligned} \tag{2-34}$$

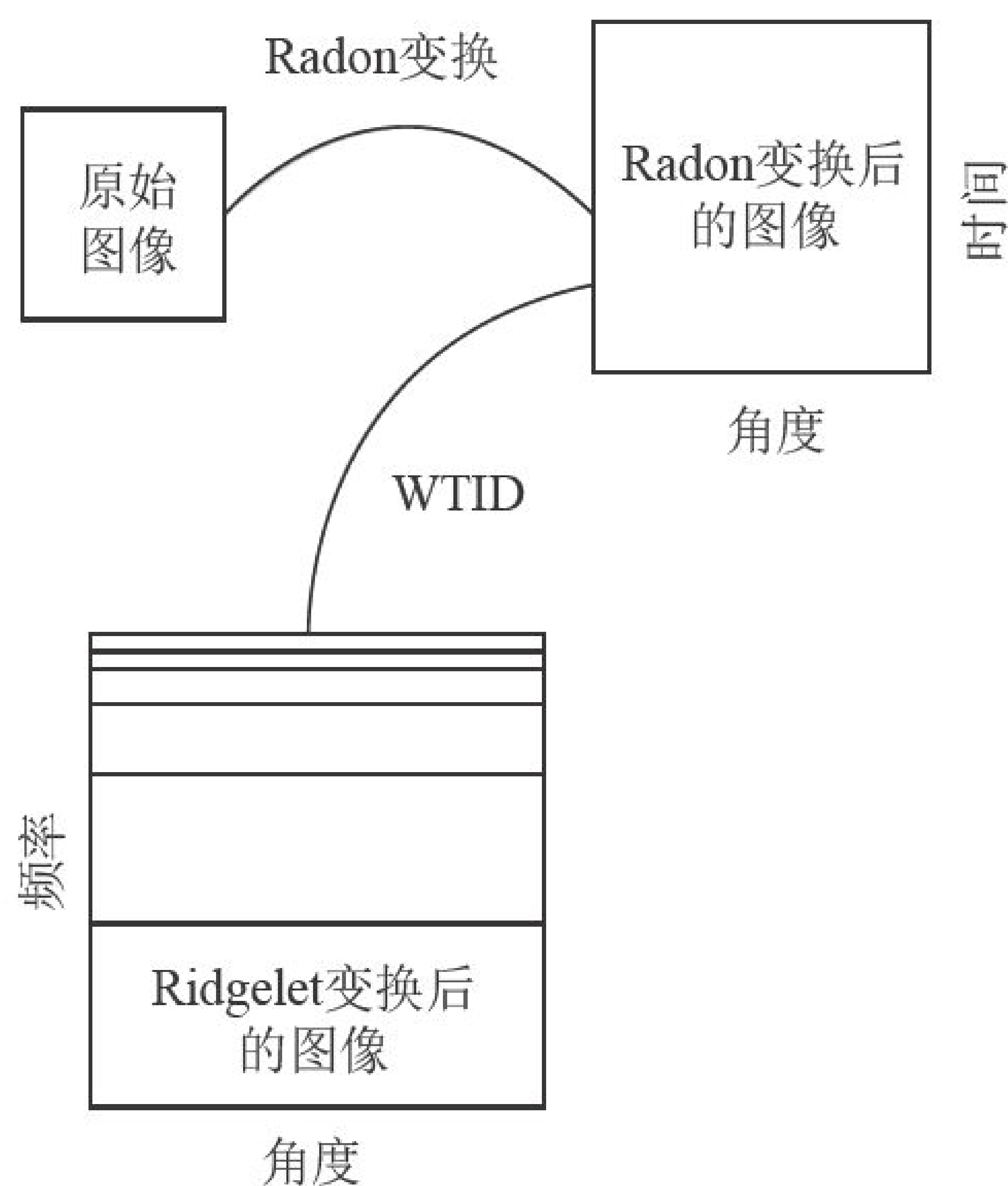


图2-18　Ridgelet变换过程

令$m=c=p=0.5$，$s=0$，其中p决定非线性度，s表示图像的饱和度；c为归一化参数，在$c>3$时保证噪声不会得到增强；m可以由噪声的标准方差$m=k_m\sigma$或者从Curvelet变换的最大系数$m=LMc$，$L<1$得到。[11][12]基于Curvelet变换的图像对比度增强过程如下所述。

(1) 估计输入图像的标准噪声方差σ。

(2) 计算图像的Curvelet系数，得到一组子带w_j，每一个子带w_j包含N_j个系数，对应于一个给定的分辨率。

(3) 计算每一个Curvelet变换子带w_j中的标准噪声方差。

(4) 针对每一个w_j计算其中的最大值M_j，然后用$y(|w_{j,k}|, \sigma_j)$和Curvelet变换的系数相乘。

(5) 由经过修正的Curvelet系数重构图像，便实现了图像的对比度增强。

3. 组合增强算法(CEA)

组合增强算法包括以下4个步骤。

(1) 对图像I进行分割，分割成平滑区域和边缘部分。平滑区域可以通过计算图像平方差来得到，其每个像素被定义为

$$V[i, j]=\sum_{p, q\in c}[I(i, j)-I(p, q)]^2 \tag{2-35}$$

这里C为以$[i, j]$为中心的11×11窗口，对于图像I中的平滑区域，图像平方差应接近于噪声方差σ^2，则可以通过阈值检测平滑像素

$$V[i, j]/\sigma^2\leqslant(1+\delta) \tag{2-36}$$

δ为一个非常小的数，其他不符合条件的像素被定义为边缘像素。

(2) 对图像I的平滑区域进行DWT增强。

(3) 对图像I的边缘部分进行Curvelet增强。

(4) 利用B、C中得到的两幅图像进行融合。设平滑区域的像素值为$I_1(x, y)$，边缘区域的像素值为$I_2(x, y)$，经过计算后得到新的像素值为$J(x, y)$，其中$x, y\in N$。采用空域中点处理方式，逐个像素点进行运算。为了方便处理，先将两幅图像划分为3×3的矩形区域，在处理过程中，设定一个融合函数，使其满足如下关系

$$J(x, y)=FS(x, y)[I_1(x, y)-I_2(x, y)]+PI_2(x, y) \tag{2-37}$$

其中，F为融合因子，规定了融合算法的步长，这里取$f=0.1$，$S(x, y)$为与位置坐标有关的系数，规定融合是图像渐变的方式。$S(x, y)$取为

$$S(x,\ y)=\sqrt{(x-R_x)^2+(y-R_y)^2} \tag{2-38}$$

其中，R_x、R_y是矩形区域的中心。

CEA原理图，如图2-19所示。

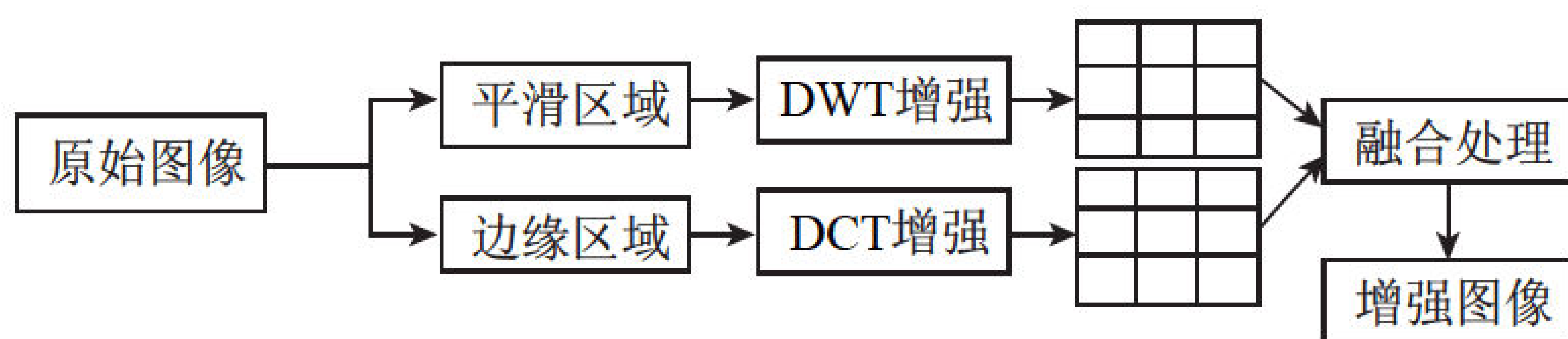


图2-19　CEA原理图

2.4　实验结果与分析

可进行融合增强结果评价的图像统计参数包括图像的交叉熵、增强前后图像的相关系数等。

(1) 交叉熵(Cross Entropy)。交叉熵可用来度量两幅图像灰度分布信息的差异，交叉熵越小，说明融合后图像与标准参考图像的差异越小，效果越好。若标准参考图像为R，融合图像为F，则交叉熵为

$$\mathrm{CE}=\sum_{i=0}^{L-1}P_{R_i}\log\frac{P_{R_i}}{p_{F_i}} \tag{2-39}$$

其中，P_{R_i}为图像R中灰度值为i像素数与图像总像素之比，P_{F_i}为图像F中灰度值为i的像素数与图像总像素数之比。

(2) 相关系数CORR。图像的相关系数反映了两幅图像的相关程度。

$$\mathrm{CORR}=\frac{\sum_{i=1}^{M}\sum_{j=1}^{N}[R(i,j)-\overline{R}][F(i,j)-\overline{F}]}{\sqrt{(\sum_{i=1}^{M}\sum_{j=1}^{N}[R(i,j)-\overline{R}])^2(\sum_{i=1}^{M}\sum_{j=1}^{N}[F(i,j)-\overline{F}])^2}} \tag{2-40}$$

其中，$\overline{R}$、$\overline{F}$分别为图像R、F的灰度平均值。实验结果如表2-3所示。

表2-3 图像增强评价指标

增强方法	交叉熵	相关系数
直方图均衡化	2.175	0.956 3
线性拉伸	2.168	0.962 4
CEA算法	2.017	0.973 8

原图像如图2-20所示。

对原图像进行直方图均衡化处理后的图像结果，如图2-21所示。



图2-20 原始图像



图2-21 直方图均衡化后的结果

对原图像进行线性拉伸的结果，如图2-22所示。

对原图像进行CEA增强后的结果，如图2-23所示。



图2-22 线性拉伸的结果



图2-23 CEA增强后的图像

边缘增强的效果用对比度改善指数CII(The Contrast Improvement Index)来衡量。CII

定义为

$$CII=C_{processed}/C_{original} \tag{2-41}$$

将图像分成3×3的小块图像，C为所有3×3小块图像对比度的平均值，对比度定义为(max−min)/(max+min)，max为该块图像灰度值的最大值，min为该块图像灰度值的最小值。计算如图2-24所示范围内的对比度改善指数，结果如表2-4所示。

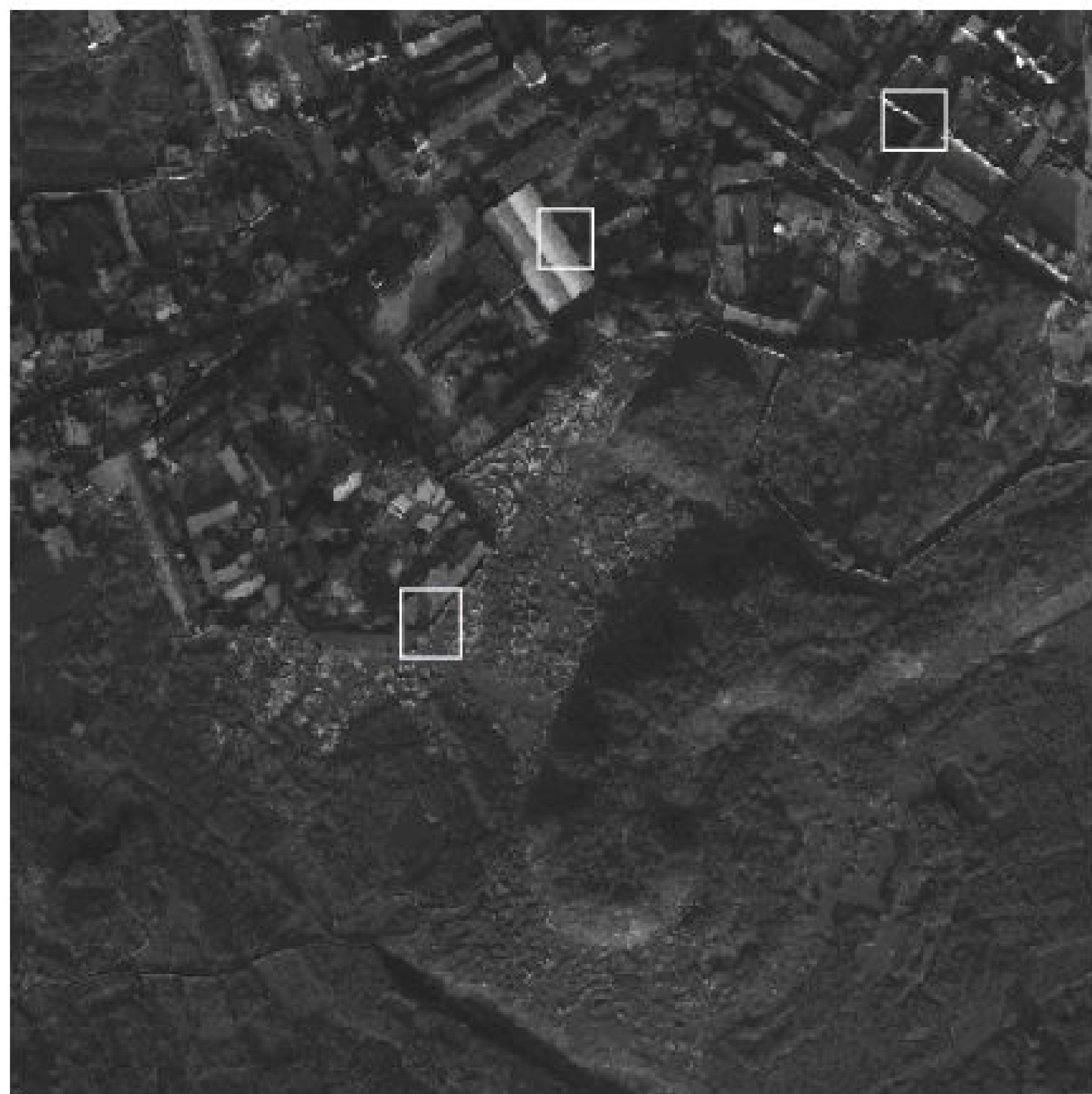


图2-24　CII计算范围

表2-4　对比度改善指数

计算范围所在位置	右上	中	左下
直方图均衡化	1.464 0	1.785 3	1.972 4
线性拉伸	1.598	1.827 4	2.014 8
CEA算法	1.892 1	2.097 4	2.759 2

2.5　结论

本章主要研究图像拼接的前期处理工作，包括图像校正和增强处理。几何校正是遥感图像处理中的重要内容，具有计算量大、耗时长的特点。为了加快几何校正的处理，针对遥感图像，我们研究并实现了一种分布存储环境下的并行几何校正算法。 在该算法中，每个处理器通过计算本地输入子图像在目标图像中的对应区域，并对该区域内的像元进行灰度值重采样，使并行计算过程中所需的数据全部为本地数据，整个计算过程无须通信，

很好地解决了数据局部性问题，大大提高了并行效率。本章对输出图像块的存储结构与局部输出区域的计算方法进行了详细讨论，并在机群系统上实现算法。利用大小不同的遥感图像进行测试，结果表明算法具有良好的并行性能，且网络带宽越高，算法的可扩展性越好。

另外，本章采用CEA增强算法，充分结合小波变换多尺度分析的能力和Curvelet变换对方向的高度敏感性，对图像分割后采用不同尺度上的系数进行不同程度的增强，以避免单一手段的增强方法在图像质量方面的不足。在去除噪声的同时，它保留了边缘细节，明显改善了图像质量，后续特征点提取的准确率也有了明显提高；增强后的图像上，图像的对比度有了明显提高，尤其将图像中一些比较模糊的区域变得清晰。它可应用于一般的低对比度图像的边缘提取，进而对图像中感兴趣的区域进行分割。实验结果表明，CEA算法与传统的增强算法相比，提高了对比度改善指数，增强效果很好。

第3章
灰度序列遥感图像拼接



通过机上图像数据压缩与传输系统，无人机将快视图无线实时下传到地面移动接收设备。对于无人机遥感监测系统获取的遥感图像，需要实现快速的实时序列图像的拼接，从而得到地物全局图像图，以方便地面专家的实时分析和决策，保证航迹规划的正确执行。快速序列实时拼接所要完成的具体任务是接收无人机传送的快视序列图像，实时完成每一帧接收图像与上一帧图像的拼接，得到地物全局图像图，使地面监控人员能迅速观察到无人机的飞行航迹，同时判断无人机图像拍摄过程中是否存在漏拍、错拍等现象，以重新调整无人机的飞行航迹，保证无人机成功完成航拍任务。如何根据无人机图像获取的特征，提高图像匹配的实时性，是实现无人机航空遥感图像拼接的关键。

由于成像时间、成像角度和成像机理不同，这些图像对同一地面特征的反应不尽相同，通过研究发现，主要相似点是道路、建筑物、植被区等的形状特征。由此可见，景物的边缘形状特征是这些图像共同具有的特征，因此，对于不同性质的图像的配准，研究其边缘特征，尤其是其中大的边缘特征，是进行精确匹配的关键。因为大的边缘是高频信息，是不同传感器共有的信息，并在不同分辨率的图像中都能保持相对稳定，都比较容易被检测。另外，采用大的边缘作为匹配特征，可以减少小的边缘的干扰，从而提高配准的精度和速度，为顺利实现拼接打下良好的基础。

本章提出一种小波变换与Canny算法相结合的遥感图像大边缘检测方法，该方法是为图像配准而设计的。该方法将小波变换多分辨率分析特性和Canny算法检测边缘思想相结合，提取遥感图像中大的边缘，确定两幅图像的重叠范围，提取具有丰富图像信息的模板，设计模板匹配搜索策略，完成重叠图像区域的粗匹配和细匹配，得到两帧图像相对的位置偏移。根据图像匹配的计算结果，完成前后两帧图像的对接，并对图像拼接过渡区域的幅度值平滑化。

3.1 图像边缘检测

图像的边缘是图像最基本的特征，是图像的一个属性区域和另一个属性区域的交接处，是区域属性发生突变的地方，是图像中不确定性最大的地方，也是图像信息最集中的地方，图像的边缘包含丰富的信息。常见的边缘检测算法有以下几种。

3.1.1　微分算子法

经典的图像边缘检测的微分算子有Roberts算子、Sobel算子、Prewitt算子、Krisch算子、拉普拉斯-高斯(LOG)算子等。在这些算子中，除了拉普拉斯-高斯算子，其他算子(如Sobel算子、Prewitt算子和Krisch算子)都是对像素点的邻域操作，即对像素点及其周围的一定邻域进行综合计算，并根据这个计算值来判断该像素是否构成边缘点，并没有考察边缘点之间的联系，梯度方向信息也没有得到充分利用，无法区分边缘和噪声，无法检测出图像中不同尺度上的边缘，对于细节丰富的图像和含噪声的图像，检测效果很差。

LOG是基于二阶导数过零点的边缘检测方法，被认为是微分法中利用平滑二阶微分检测图像边缘最成功的一种算子。LOG是通过寻找图像灰度值中二阶微分中的过零点来检测边缘点。其原理是灰度缓变形成的边缘经过微分算子形成一个单峰函数，峰值位置对应边缘点；对单峰函数进行微分，则峰值处的微分值为0，峰值两侧符号相反，而原先的极值点对应二阶微分中的过零点，通过检测过零点即可将图像的边缘提取出来。

在实际中，为了去除噪声的影响，首先要使用二维高斯函数对图像进行去噪音和平滑处理，然后对滤波后的图像求二阶导数。

$$H(x,\ y)=\nabla^2 G(x,\ y)\times f(x,\ y) \tag{3-1}$$

其中

$$\nabla^2 G(x,\ y)=\frac{1}{\pi\sigma^4}[1-(\frac{x^2+y^2}{2\sigma^2})]\cdot\exp(\frac{x^2+y^2}{2\sigma^2})$$

使用LOG算子对图像进行处理后，就可以提取零交叉点确定边缘。LOG算法的优势是去除噪音的性能好，提取边缘的精度高，可以达到像素级；但缺点也很明显，一是容易产生伪边缘，二是邻近的边缘之间互相干扰，提取的边缘可能有误。

3.1.2　Canny算法

Canny算法认为一个优良的边缘检测算子应满足以下3个判断准则。

(1) 好的检测性能。不漏检真实边缘，也不把非边缘点作为边缘点检出，使输出的信噪比最大。

(2) 好的定位性能。检测到的边缘点与实际边缘点位置之间的偏差最小。

(3) 唯一性。对一个边缘只有一个响应。

根据以上三个准则，Canny算法推导出最优边缘检测算子的一个近似实现是边界点位于图像被高斯函数平滑后的梯度幅度的极大值点。该算法主要包括以下几个步骤。

(1) 平滑图像。用一维高斯函数分别按行和列对原始图像进行平滑除噪，得到平滑图像。这一步实际上是一个低通滤波过程，用于消除空间尺度小于高斯空间系数的图像灰度变化。

(2) 计算梯度的幅值和方向。对平滑图像使用如图3-1所示的2×2大小模板作为对x和y方向偏微分的一阶近似来计算其梯度的大小和方向。

−1/2	1/2
−1/2	1/2

$g(x)$

1/2	1/2
−1/2	−1/2

$g(y)$

图3-1　图像平滑模板

(3) 对梯度幅值进行非极大值抑制。为了精确定位边缘，必须细化梯度幅值图像中的屋脊带，只保留幅值局部变化最大的点。

(4) 检测和连接边缘。对非极大值抑制后形成的梯度幅值图采用双阈值算法对边缘点进行检测和连接，形成边缘。

Canny算法在边缘检测方面获得了良好的效果，很快就成为评价其他边缘检测方法的标准。然而，Canny算法在实际应用中也存在一些问题。比如，Canny算法的噪声平滑能力与边缘定位能力相矛盾。

3.1.3　数学形态学在边缘检测中的应用

数学形态学是建立在严格数学理论上的一门新兴学科[156]，它已从二值形态学发展到灰度形态学，是一种新型数字图像处理方法和理论。它的基本思想是用具有一定形态的结构元素去提取图像中对应的形状，以达到对图像分析和目标识别的目的。它可以用来解决抑制噪声、特征提取、边缘检测等图像处理问题。[177][178]数学形态学中二值图像的形态变换的基本运算是腐蚀和膨胀，腐蚀运算可以去掉图像中比结构元素小的亮细节，膨胀运算可以去掉比结构元素小的暗细节。先腐蚀后膨胀的过程称为开启运算，开启运算可以消除

比结构元素小的亮细节，而保持图像整体灰度和大的亮区域基本不受影响。先膨胀后腐蚀的过程称为闭合运算，具有填充物体内细小空洞，连接临近物体和平滑边界的作用。该算法简单，适用于并行处理，且易于硬件实现，适用于对二值图像进行边缘提取。[178]对于在不同的应用场合，结构元素的选择及其相应的处理算法是不一样的，结构元素的大小、形状选择合适与否，将直接影响图像的形态处理结果。有学者[179][180]构造多结构元的边缘检测算子，用多结构元进行边缘检测，不仅具有较好的边缘定位能力，也具有很好的噪声平滑能力。形态学边缘检测算子中的结构元素不仅有形状，而且有大小(如3×3、5×5的窗口)，因此可以用多尺度的结构元素对图像进行多尺度的边缘检测。有学者[181][182]利用不同的结构元素在不同尺度上进行边缘检测，由于大尺度的结构元素检测的边缘较粗，但去噪声能力强，小尺度的结构元素去噪声能力弱，但能很好地检测到边缘细节，将两者结合可得到较理想的边缘图像。Soille[183]将灰度形态学梯度与阈值相结合用于边缘检测，可较好地检测出被噪声污染的图像中的边缘。

从以上分析中可以看出，现有的边缘检测算法都是用于提取一般的边缘，而在遥感图像配准中对检测出的边缘有特定的要求，它要求检测出来的边缘是一种比较稳定、比较大的边缘，在不同时间、不同角度、不同传感器所呈现的图像中都能被检测出来，同时要求检测出来的边缘特征分布比较均匀。所以，这些已有的方法不能很好地解决这个问题。为了满足遥感图像配准的要求，本书提出了一种结合小波变换和Canny算法的边缘提取算法。该算法将Canny算法的非极大值抑制、双阈值算法与小波变换多尺度分析思想相结合，利用Canny算法在高分辨率图像的精确定位性与小波变换后的低分辨率图像的抗噪声能力相结合，从而能够消除噪声的干扰，提取更加完整的边缘。同时，该算法采用自适应阈值的方法，能减少由于噪声和图像对比度差异所引起的边缘漏检情况的发生。

3.1.4　基于小波变换的边缘检测方法

取一个二维光滑函数$\theta(x，y)$，其满足以下条件

$$\begin{cases} \theta(x，y)>0 \\ \iint_{L^2(R^2)} \theta(x，y)\mathrm{d}x\mathrm{d}y=C \end{cases} \tag{3-2}$$

把它在x、y方向的一阶偏导作为两个基本小波，则

$$\psi_s^{(1)}(x,\ y)=\frac{\partial_s\theta(x,\ y)}{\partial x}$$
$$\psi_s^{(2)}(x,\ y)=\frac{\partial_s\theta(x,\ y)}{\partial y} \tag{3-3}$$

则$f(x,\ y)\in L^2(R^2)$的小波变换定义为

$$\begin{cases} W_s^1 f(x)=f*\psi_s^{(1)}(x) \\ W_s^2 f(x)=f*\psi_s^{(2)}(x) \end{cases} \tag{3-4}$$

若以二进制的小波变换表示，即s取2^j，则有如下两个分别沿x、y方向的分量

$$\begin{cases} W^1 f(x,\ y)=f*\psi_{2^j}^{(1)}(x,\ y)=2^j\dfrac{\partial}{\partial x}(f*\theta_2\ )(x,\ y) \\ W_s^2 f(x,\ y)=f*\psi_{2^j}^{(2)}(x,\ y)=2^j\dfrac{\partial}{\partial x}(f*\theta_{2j})(x,\ y) \end{cases} \tag{3-5}$$

式(3-5)可记为矢量形式

$$\begin{bmatrix} W_{2^j}^1 f(x,\ y) \\ W_{2^j}^2 f(x,\ y) \end{bmatrix}=2^j\begin{bmatrix} \dfrac{\partial}{\partial x}(f*\theta_{2j})(x,\ y) \\ \dfrac{\partial}{\partial y}(f*\theta_{2j})(x,\ y) \end{bmatrix}=2^j\vec{\nabla}(f*\theta_{2j})(x,\ y) \tag{3-6}$$

由此可以看出，小波变换的两个分量与$f(x,\ y)$被$\theta(x,\ y)$所平滑的梯度向量的坐标成正比。$f(x,\ y)$在尺度s的边缘点就是这一梯度向量模的最大值点，而这一梯度向量的方向正是$f(x,\ y)$的偏导数变化最迅速的方向。在每个尺度$s=2^j$上，梯度幅值和梯度方向分别为

$$M_{2^j}f(x,\ y)=\sqrt{\left|W_{2^j}^1 f(x,\ y)\right|^2+\left|W_{2^j}^2 f(x,\ y)\right|^2}$$
$$A_{2^j}f(x,\ y)=\tan\left(\left|W_{2^j}^1 f(x,\ y)\right|/\left|W_{2^j}^2 f(x,\ y)\right|\right) \tag{3-7}$$

因此，可以在图像平面上沿着梯度方向$A_{2^j}f(x,\ y)$检测梯度幅值$M_{2^j}f(x,\ y)$的极大值点，然后根据选取阈值的某种方法确定一个阈值T，以取舍该极大值点，再进行二值化就得到图像的边缘。

3.2 小波变换与Canny算法相结合的边缘检测方法

本书提出的小波变换与Canny算法相结合的自适应边缘检测算法的主要思路是：将遥

感图像先利用小波变换进行多分辨率分解，生成不同分辨率的图像，再对每个分辨率的图像计算梯度的幅值、方向和对梯度幅值进行非极大值抑制，然后结合多层梯度幅值信息采用双阈值算法对边缘点进行检测和连接，形成边缘。该算法的主要步骤如下所述。

1. 生成多尺度梯度图像

先应用离散小波变换将原始图像进行多尺度分解。对原始图像通过2D离散小波分解可以得到4个子图像，一个低频信号LL_1和三个方向的高频分量信号LH_1、HL_1、HH_1。在第一层分解时，不进行下采样(Down Sample)，所以得到的4个子图像与原始图像的大小相同。低频信号LL_1，又可以进一步分解成4个子带，可以根据需要继续往下分解，一般分解到第三层、第四层。分解完成后分别计算每一层图像的梯度幅值$M_{2^j}f$，生成相应的梯度幅值图像，并计算梯度方向$A_{2^j}f$。

2. 对梯度进行非极大值抑制

梯度幅值图像阵列$M(i，j)$的值越大，其对应的图像梯度值也越大，但这不足以确定边缘，为了确定边缘，必须细化幅值图像中的屋脊带，这样才会生成细化的边缘。非极大值抑制通过抑制梯度方向上所有非屋脊峰值的幅值来细化$M(i，j)$中的梯度幅值屋脊。这个过程具体如下：首先将梯度方向$A(i，j)$变成属于如图3-2所示的4个区之一，再使用3×3大小，包含8个方向的邻域对梯度幅值阵列$M(i，j)$的所有像素沿梯度方向进行梯度幅值的插值。在每一个点上，邻域的中心像素与沿梯度方向的2个梯度幅值的插值进行比较，如果邻域中心点的幅值不比梯度方向上的2个插值结果大，则将该点幅值$M(i，j)$标记为0。经过这一过程处理后宽屋脊带细化成只有一个像素点宽，在非极大值抑制过程中，保留了屋脊的高度值。对每一层的梯度图像都进行非极大值抑制，生成相应的非极大值抑制幅值图。

3	2	1
0	×	0
1	2	3

图3-2　扇区示意图

3. 阈值化和边缘连接

阈值化是对非极大值抑制幅值图梯度图像进行双阈值处理生成二值化图像。高、低闭值Th和Tl，设定方法是一种自适应动态阈值方法。对第一层生成的非极大值抑制幅值图梯度图像取两次阈值Th和Tl。首先将梯度值小于Tl的像素的梯度设为0，得到二值化图像I_{11}；然后将梯度值大于Th的像素的梯度设为1，得到二值化图像I_{12}。对其他层生成的非极大值

抑制幅值图梯度图像取一次阈值Tl，生成相应的二值化图像I_k。图像I_{12}阈值较高，噪声较少，但会造成边缘信息损失；而图像I_{11}阈值较低，保留了较多信息，图像I_k是根据低分辨率图像生成的，受噪声影响小。因而可以以图像I_{12}为基础，以图像I_{11}和I_k为补充来连接图像的边缘。

根据多层边缘图连接边缘的过程如下：首先在图像I_{12}中扫描，当遇到一个非零的像素P时，跟踪以P为起始点的轮廓线，直到该线的终点Q；然后根据以下准则判断终点Q的8邻近区域是否存在与终点Q相连的边缘点。

其中，O是8邻近区域中的点，O'是O点在I_k中的对应点。如果$C(o)$=1，表示O点是与Q点相连的边缘点，将其纳入图像I_{12}中，作为边缘点，然后继续跟踪，在图像I_{12}中重复继续寻找与跟踪以P为开始点的轮廓线，这样循环下去直到无法继续为止，包含P的轮廓线的连接已经完成，可标记为“已访问过”，然后依次可以重复寻找图像中的每一条轮廓线，直到在图像I_{12}中再也找不到新的轮廓线为止。

$$C(o)=\begin{cases}1 & I_{11}(0)=1\\ 1 & I_k(o')=1\text{且}I_{11}(0)\text{是局部梯度最大点。}\\ 0 & \text{其他}\end{cases}$$

采用根据多层边缘图连接边缘的方法的优点是利用了在高分辨率下边缘点的定位精确性，同时又利用了粗分辨率图像的抗噪性，使连接的边缘更加完整。

3.2.1　小波基选择

小波的种类很多，如Morlet小波、Harr小波、Marr小波、Daubechies紧支正交小波以及样条小波等，也可以自己构造小波。每种小波都具有各自不同的特点和应用领域，因此应根据要解决问题的特点来选取。一般而言，在图像边缘检测应用中，小波应满足以下准则。

准则一：紧支性。若函数$\psi(x)$在区间[a，b]外恒为零，则称该函数在这个区间上紧支。紧支性是小波的重要性质，支集越小的小波，局部化能力越强，不需要做截断，也就无截断误差，所以精度比较高，在图像的边缘检测中，紧支小波基是首选。

准则二：对称性。对称性能避免信号在分解与重构中失真，基于人类的视觉系统对边缘附近对称的量化误差较非对称误差更不敏感。利用小波的对称性可以避免边界失真，同

时若小波具有对称性，则在频域上为线性相位，没有相位畸变而且计算方便。

准则三：适度的正则性。正则性一般用来刻画函数的光滑程度，正则性越高，函数越光滑，在频域中能量越集中。对非紧支集小波，小波滤波器越光滑则收敛越快，这样小波系数衰减就越快，截断误差就越小。尺度函数的正则性与小波函数的消失矩成对应关系，高阶正则性的尺度函数对应高阶消失矩的小波函数。但是正则性与计算效率常常是一对矛盾，正则性越高，小波消失矩越大，计算越复杂。

研究[159]表明，从时频局部分析的角度在噪声抑制能力和边缘提取方面三次B样条函数是渐进最优的。因此选用三次B样条函数作为平滑函数，其一阶导数作为小波函数。这是一个关于原点奇对称的小波，不会产生时间(或空间)的偏移，并且是紧支撑和连续可微的。

3.2.2　自适应确定双阈值

传统的Canny算法采用对整幅图像使用固定的高、低阈值设定方法，无法顾及图像中的局部特征信息，一方面无法消除局部噪声干扰，另一方面会丢失灰度值变化缓慢的局部边缘，导致目标物体的轮廓边缘不连续或者检测不到一些目标。对于一幅具体图像，传统的Canny算法检测边缘时最佳的高、低阈值完全依赖人工获得，自动化程度低。本书采用一种自适应动态阈值方法，将整幅图像分割为若干子图像，为了使轮廓连续，可以令子图像之间有一定的重叠区域，再根据非极大值抑制后的结果自适应地设定各子图像的高、低阈值。

由于一般图像中只有少量的像素是边缘，所以图像中非边缘所占比例远大于边缘所占比例，因此梯度幅值分布一般是单峰分布，并可以认为单峰分布的峰顶对应的像素集合一定是非边缘像素集合。将梯度幅值中拥有最多像素的梯度称为像素最值梯度H_{max}，计算子图像内全部像素的梯度相对于像素最值梯度H_{max}的方差，称为像素最值梯度方差σ_{max}。

$$\sigma_{max}=\sqrt{\sum_{i=0}^{N}(H_i-H_{max})^2/N} \tag{3-8}$$

其中，N为子图像的像素总数。

高阈值T_h必须在梯度直方图中非边缘区域以外选取，否则将给最终结果带来很多的假边缘噪声。本书所采用的方法利用像素最值梯度H_{max}和像素最值梯度方差σ_{max}自适应地设定高阈值T_h的门限，像素最值梯度σ_{max}反映了非边缘区域在梯度直方图中分布的中心位

置，而像素最值梯度方差σ_{max}则反映了梯度直方图中梯度分布相对于像素最值梯度σ_{max}的离散程度，也就是相对于非边缘区域的离散程度，我们近似认为如果高闭值Th大于像素最值梯度H_{max}一定倍数的最值梯度方差σ_{max}时，就可以认为Th在非边缘区域以外，这样可以很好地防止假边缘在轮廓图中出现，因此可以用像素最值梯度H_{max}和像素最值梯度方差σ_{max}计算出非边缘区域的范围，从而确定高阈值Th的门限，然后根据高阈值Th的门限确定低阈值Tl的门限。计算Th和Tl的公式为

$$\begin{aligned} T_h &= H_{max} + \beta * \sigma_{max} \\ T_l &= k * T_h \end{aligned} \tag{3-9}$$

其中，β是一个调整因子，一般在2至5之间，K是高、低阈值间的一个比例系数，一般为0.4左右。采用自适应动态阈值方法设定阈值，可以减少因局部图像灰度变化差异大小不同而引起目标物体的轮廓边缘不连续或者检测不到一些目标的情况发生，同时也可以在一定程度上提高提取边缘的自动化程度。

为了验证小波变换与Canny算法相结合的边缘检测方法的有效性，分别用本书提出的方法、Canny算法和Log算法对一幅遥感影像进行边缘检测。原图如图3-3所示，图3-4为本书提出算法的提取结果，图3-5为Canny算法的提取结果，图3-6为Log算法的提取结果。

可以看出，本书提出的算法能够得到比较完整的边缘，能减少因亮度不一致而导致漏检边缘的情况发生，并且能够忽略小的边缘，能够满足图像配准的要求。

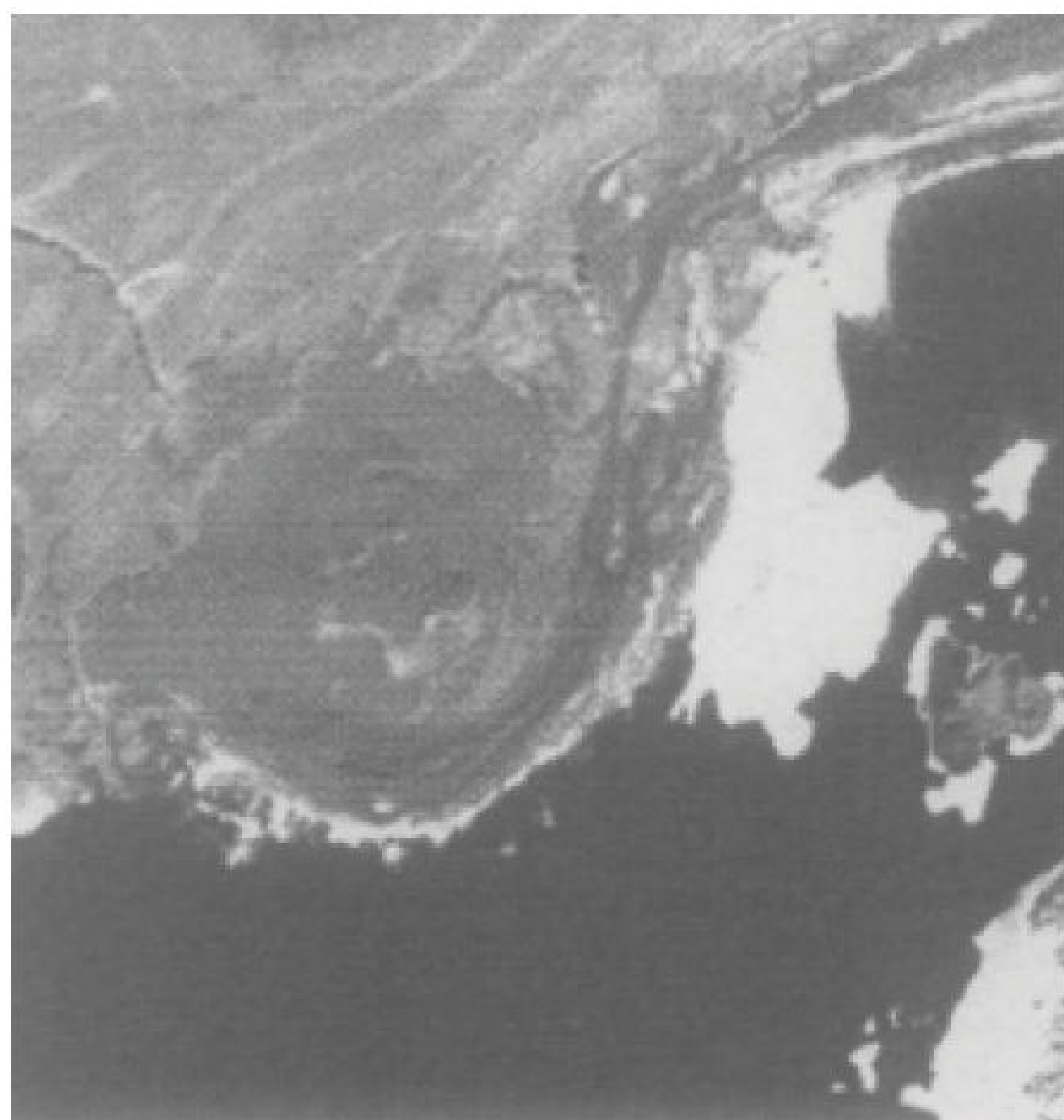


图3-3　原图

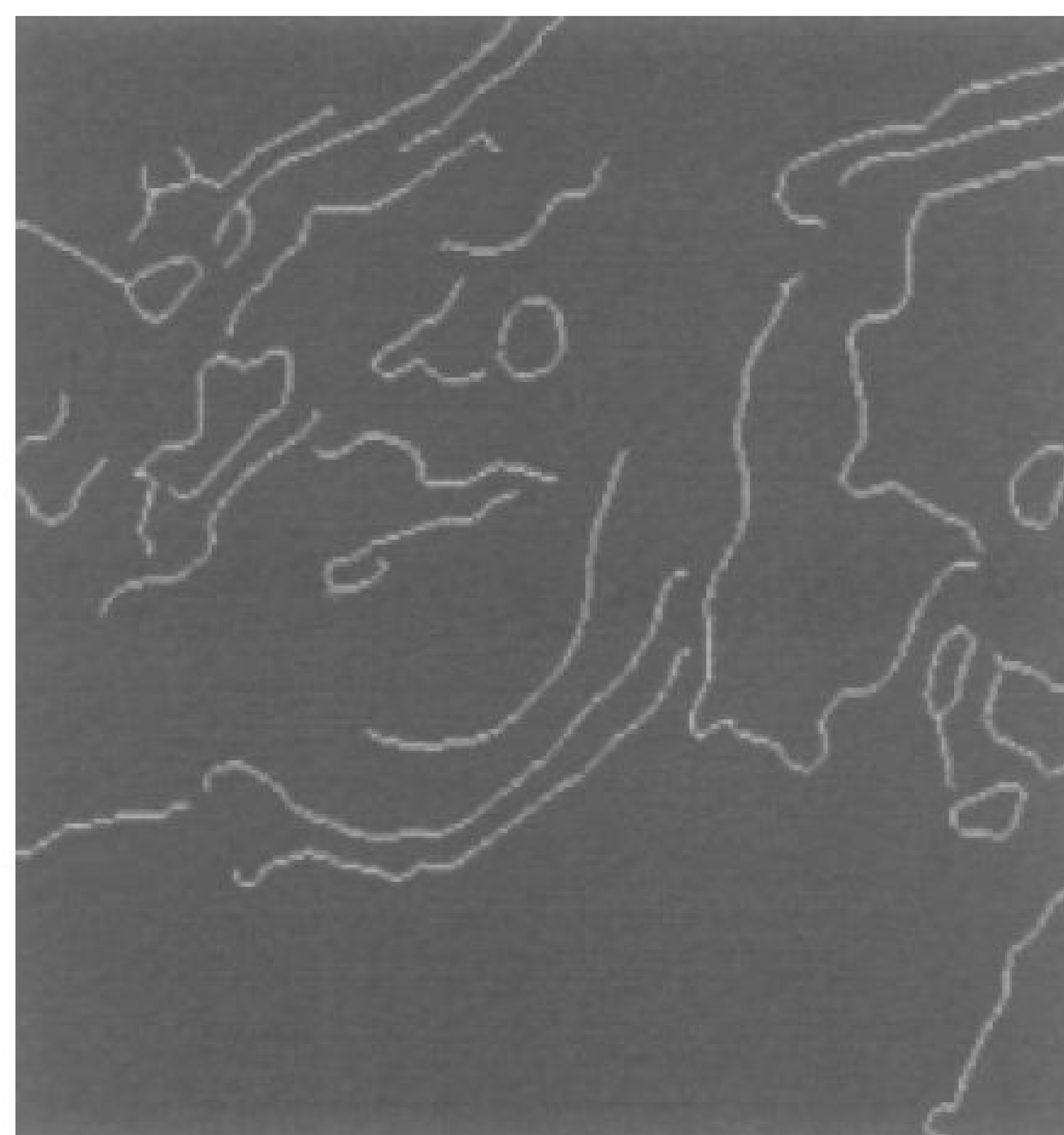
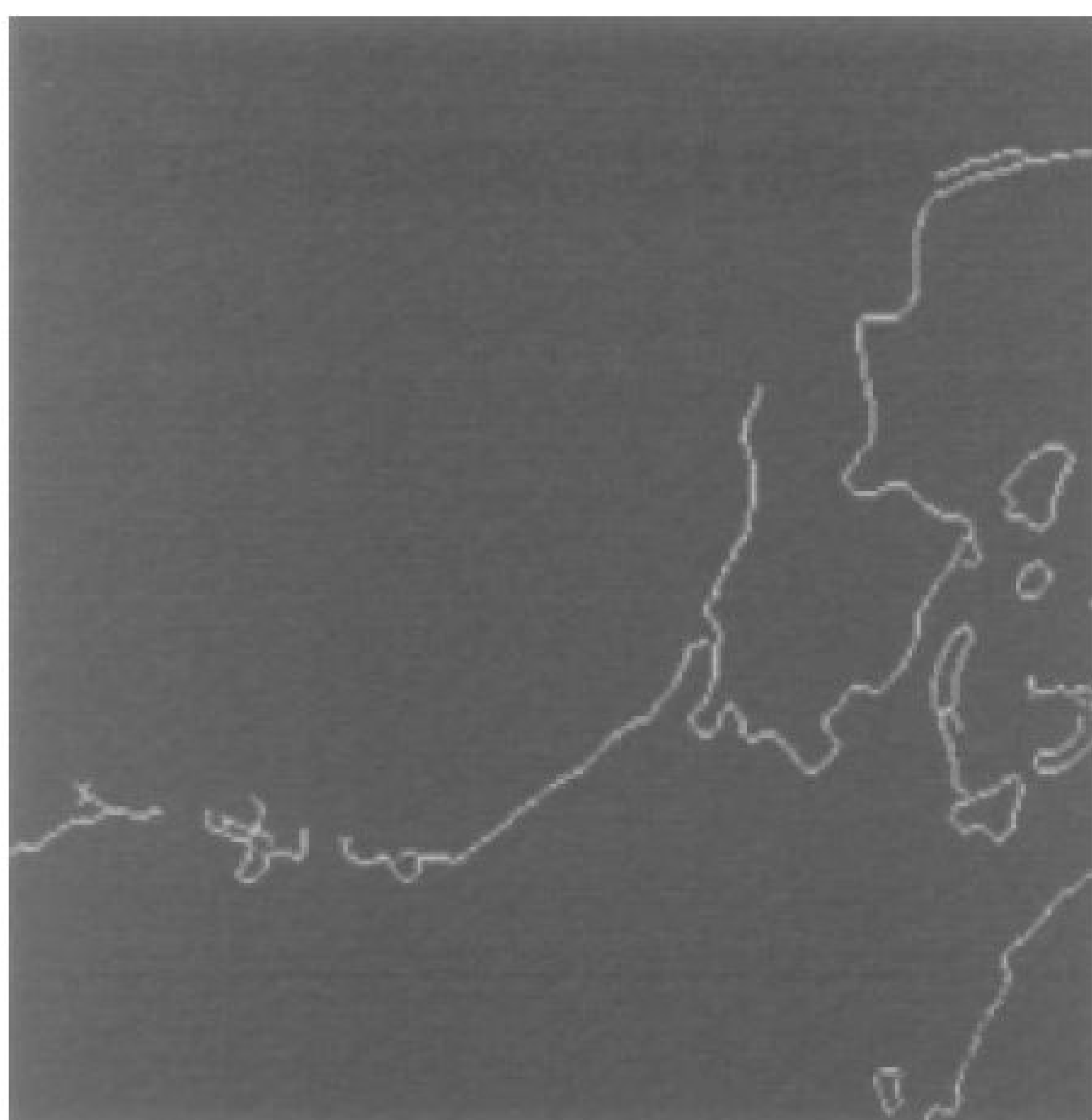


图3-4　本书提出算法的提取结果


图3-5　Canny算法的提取结果

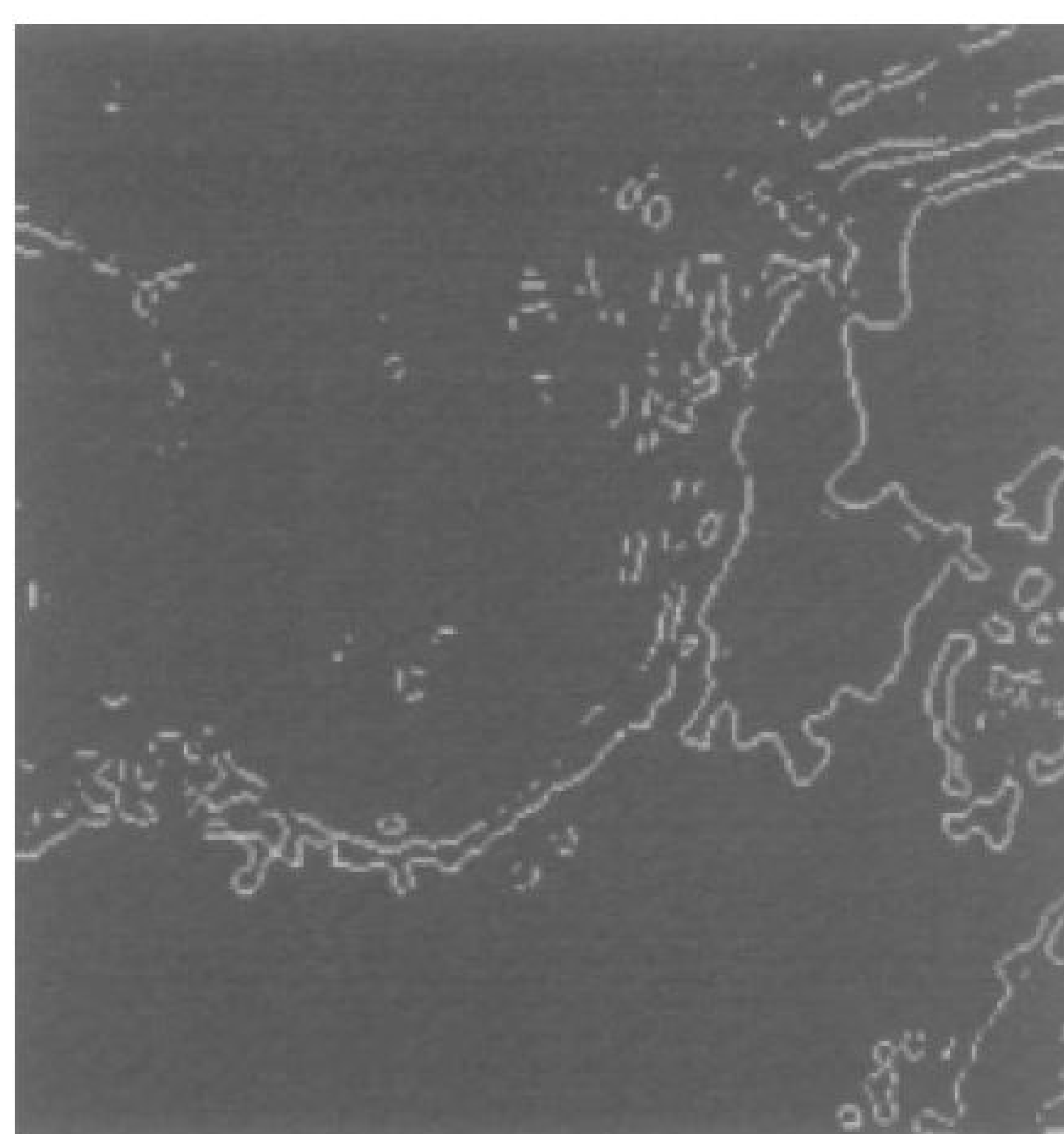

图3-6　Log算法的提取结果

图3-3　边缘检测结果

3.3　图像匹配

3.3.1　匹配搜索区域计算

首先，通过一些数据对两幅图像的重叠区域进行估计，这些数据包括无人飞机的位置、姿势和飞行速度。假设向东的速度为v_e，向北的速度为v_n，采样的时间间距为T_c，则当时与正东面的位移偏移量S_e和与正北面的位移偏移量S_n，可以进行如下的估计

$$S_e=v_e*T_c \tag{3-10}$$

$$S_n=v_n*T_c \tag{3-11}$$

用I_1和I_2分别表示两幅连续的图像，它们之间存在相对的偏移量和重叠区域。第一幅图像的行数和列数分别由M_1和N_1表示，第二幅图像的行数和列数分别由M_2和N_2表示。在图3-7中，暗灰色区域A_1表示两幅图像的重叠区域，A_2表示在重叠区域中选取的一个小区域。

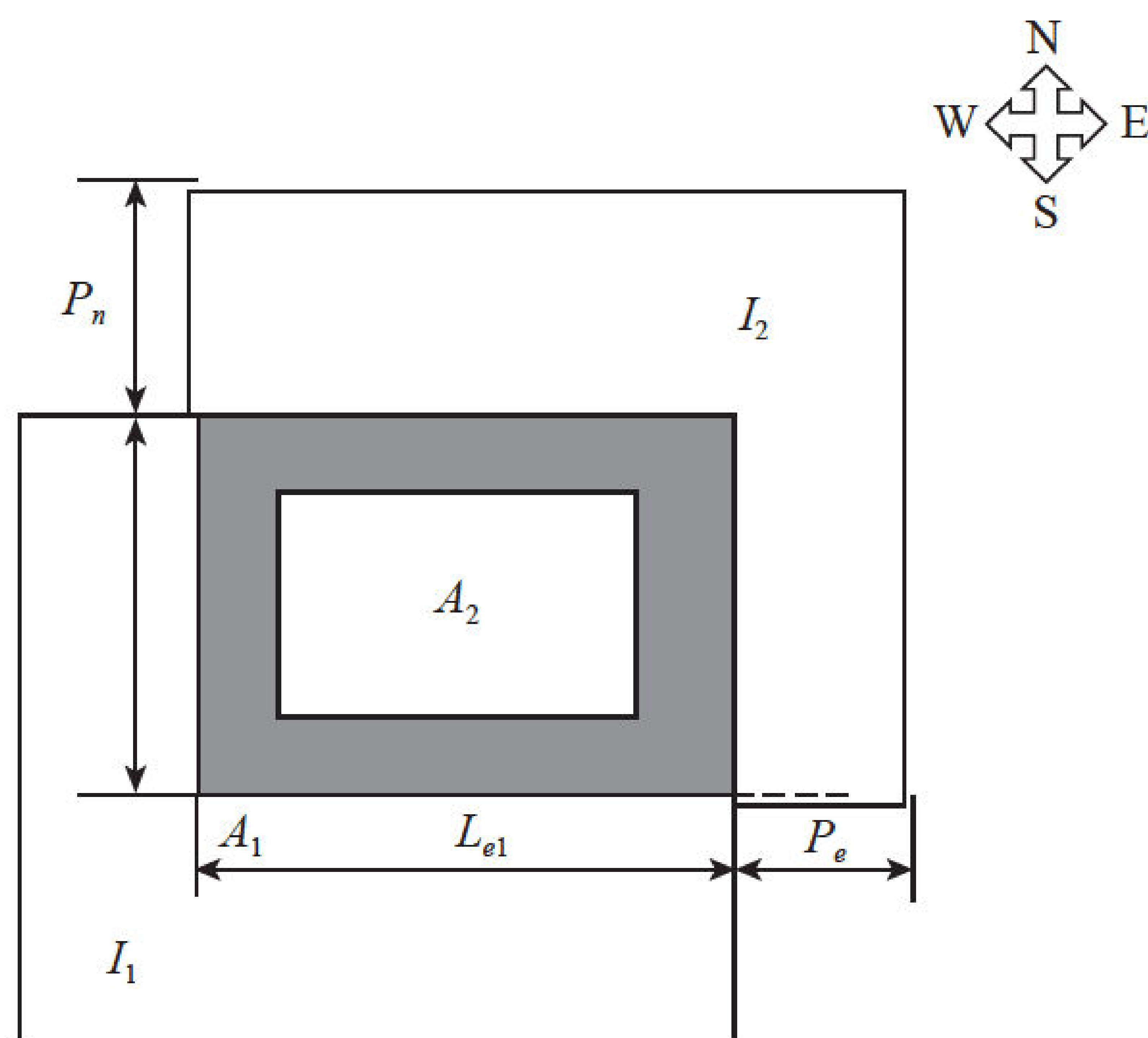


图3-7　两幅图像的重叠区域

首先对两幅图像重叠区域的范围进行估计。假设无人飞机的飞行高度为h，无人机上相机的焦距为F，每幅图像中像素的宽度为L，则地面仪器的分辨力r可以按照以下公式计算

$$r = L\frac{h}{F} \tag{3-12}$$

在T_c时间内，图像东面的像素偏移量P_e和图像北面的像素偏移量P_n分别计算为

$$P_e = {S_e}/{r} = S_e * F/(L*h) \tag{3-13}$$

$$P_n = {S_n}/{r} = S_n * F/(L*h) \tag{3-14}$$

因此，图3-7中重叠区域A_1的大小可以计算为

$$L_{e1}=M_1\square p_e \tag{3-15}$$

$$L_{n1}=N_1\square p_n \tag{3-16}$$

其中将A_1中的行数定义为L_{n1}，列数定义为L_{e1}。

为了减少图像匹配中的计算量，特征模板的选取被限定在A_2中，A_2的行数和列数分别被定义为L_{n2}和L_{e2}，它们分别为L_{e2}=L_{e1}*0.8，L_{n2}=L_{n1}*0.8。继重叠区域A_1确定之后，特征模板A_2已经被确定，可以进行下一个步骤。正因为区域A_1和A_2被限制在一个相对较小的图像范围内，所以肯定能减少在图像匹配过程中的计算量并提高其可靠性。

3.3.2　特征模板提取

对图像匹配而言，在前一幅图像中的特征模板选取是非常重要的。一些被处理的图像可能有比较光滑的区域，在那些图像中亮度的变化非常微小。如果匹配模板是在光滑的区域中选取的，那么不可避免会出现错误拼接的情况。图像匹配中包含丰富信息的特征模板是根据一定的方法选取的。建立在边缘图像统计数据基础上来选取特征模板能提高图像匹配的可靠性。

特征模板应该含有丰富的图像信息，从而使图像匹配更加容易和准确。例如在图3-7中，A_2即在重叠区域中选取的特征模板。根据图像匹配中对精确性的要求，假设特征模板的行数是M_t，列数是N_t。利用搜索步长为10，从左到右、从上到下的顺序搜索特征模板。sum(i, j)表示在(i, j)位置与模板具有相同尺寸大小的图像区域$T^{(i,\ j)}$中所有像素值的总和，计算公式为

$$\text{sum}(i,\ j)=\frac{\sum_{x=1}^{x=M_t}\sum_{y=1}^{y=N_t}T_{(x,\ y)}^{(i,\ j)}}{M_t * N_t} \tag{3-17}$$

其中$T_{(x,y)}^{(i,j)}$表示$T^{(i,\ j)}$在(x，y)坐标轴上二进制像素的值(0或者1)，其中$1<i<L_{n2}$□M_t+1，$1<j<L_{e2}$□N_t+1。假设X_g和Y_g为特征模板的灰度中心在X坐标和Y坐标上的值。符合如下情况的第一幅图像区域被确定为特征模板：

$$0.48<\text{sum}(i,\ j)<0.52 \tag{3-18}$$

$$0.4<X_g<0.6 \tag{3-19}$$

$$0.4<Y_g<0.6 \tag{3-20}$$

3.3.3　匹配计算

在获取了特征模板的基础上，可以对边缘图像进行匹配计算。本设计中的匹配计算分成两个匹配过程，即粗匹配和精匹配。其中，对边缘图像的匹配过程称为粗匹配，目的是在较短的时间内获取一些可能的匹配位置。粗匹配过程中，图像匹配中的搜索区域被限定在图3-7所示的A_1区域的图像，特征模板将会在A_1中按从上到下、从左到右的顺序移动，并进行匹配计算。匹配计算中，特征模板将会通过异或(XOR)逻辑运算对搜索区域中的每一

个边缘图像区域进行检测，这些图像区域与特征模板的尺寸大小相同。

$$\mathrm{sum}_{(u,v)}=\sum(T\otimes TB^{(u,v)}) \tag{3-21}$$

其中，T表示特征模板的二值化图像矩阵，$TB^{(u,\ v)}$表示中心在二进制搜索图像中$(u,\ v)$位置的二值化图像，⊗表示对矩阵进行异或运算，∑表示异或运算后所得结果矩阵的所有像素值的总和。

异或运算大大提高了图像匹配中计算的速度。在所有经过异或运算后所得的结果中，按公式(3-22)取前三个最小值，作为粗匹配后得到的可能位置$(u_0,\ v_0)$。

$$(u_0,\ v_0)=\min[\mathrm{sum}_{(u,v)}] \tag{3-22}$$

模板相关匹配又称基于区域的拼接方法，一般是直接或先对图像进行一定变换之后，利用图像中的像素建立模板并进行拼接，而其中模板的选取一般都不会很复杂。比较简单的方法是利用整幅图像的像素灰度为模板，但这种模板的选取往往为模板的匹配带来很大的计算量。如果为了减少匹配的计算量，只选用图像中一列图像作为模板进行匹配，那么拼接的精度往往不高，因此一般都会选取图像中某个部分区域中的像素作为模板。另外，基于傅立叶变换的相位相关法也是一种典型的算法，这种方法首先对图像进行傅立叶变换，通过频域内的相位差峰值找到图像的重叠位置，从而进行图像拼接。[12]灰度图像的空域模板匹配算法又可分为基于图像的几何特征和基于像素灰度值两大类。前者适用于以单一明确的目标检测为目的的情况，而且目标的几何特征提取(如跟踪描绘目标的边缘轮廓或提取角点等)是一个难以解决的问题，并将耗费大量的计算时间，不能满足实时系统的要求。本书使用的算法属于后者，即基于像素灰度值的匹配。

对搜索窗口和模板直接求归一化的相关值(Normalized Correlation，NC)，是一种很有效的匹配算法。设待匹配的图像区域为F，特征模板为G，则

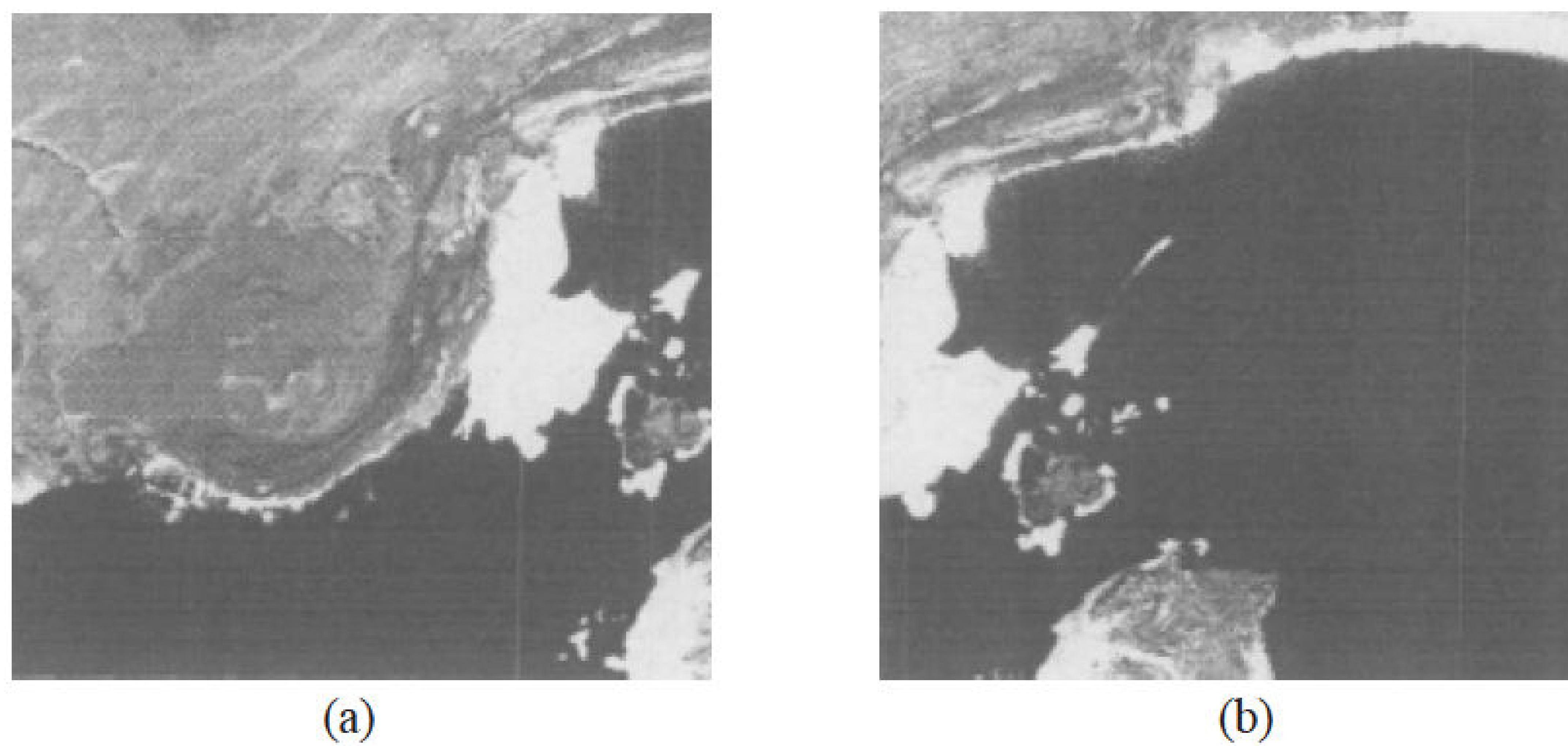
$$\mathrm{NC}=\frac{\sum\left[\left|F(x,\ y)-E(F)\right|\times\left|G(x,\ y)-E(G)\right|\right]}{\sqrt{\sum[F(x,\ y)-E(F)]^2\sum[G(x,\ y)-E(G)]^2}} \tag{3-23}$$

式中，$E(G)$表示对模板和原图的搜索窗口求数学期望，即灰度值的平均值。NC算法具有很高的准确性、适应性，对图像灰度值的线性变换具有“免疫性”，即所求的NC值不受灰度值的线性变换的影响。[17]以一个像素点为例，其过程如下：

在原图中取一个像素点记为p_0，在搜索区域中选取三个二值化图像匹配所得到的可能匹配位置，设为p_1、p_2、p_3。分别以这三个像素点为中心，选取三个5×5区域，然后分别在这三个区域中，利用公式(3-23)计算相关值，并取出每个区域中的最大相关值分别记为

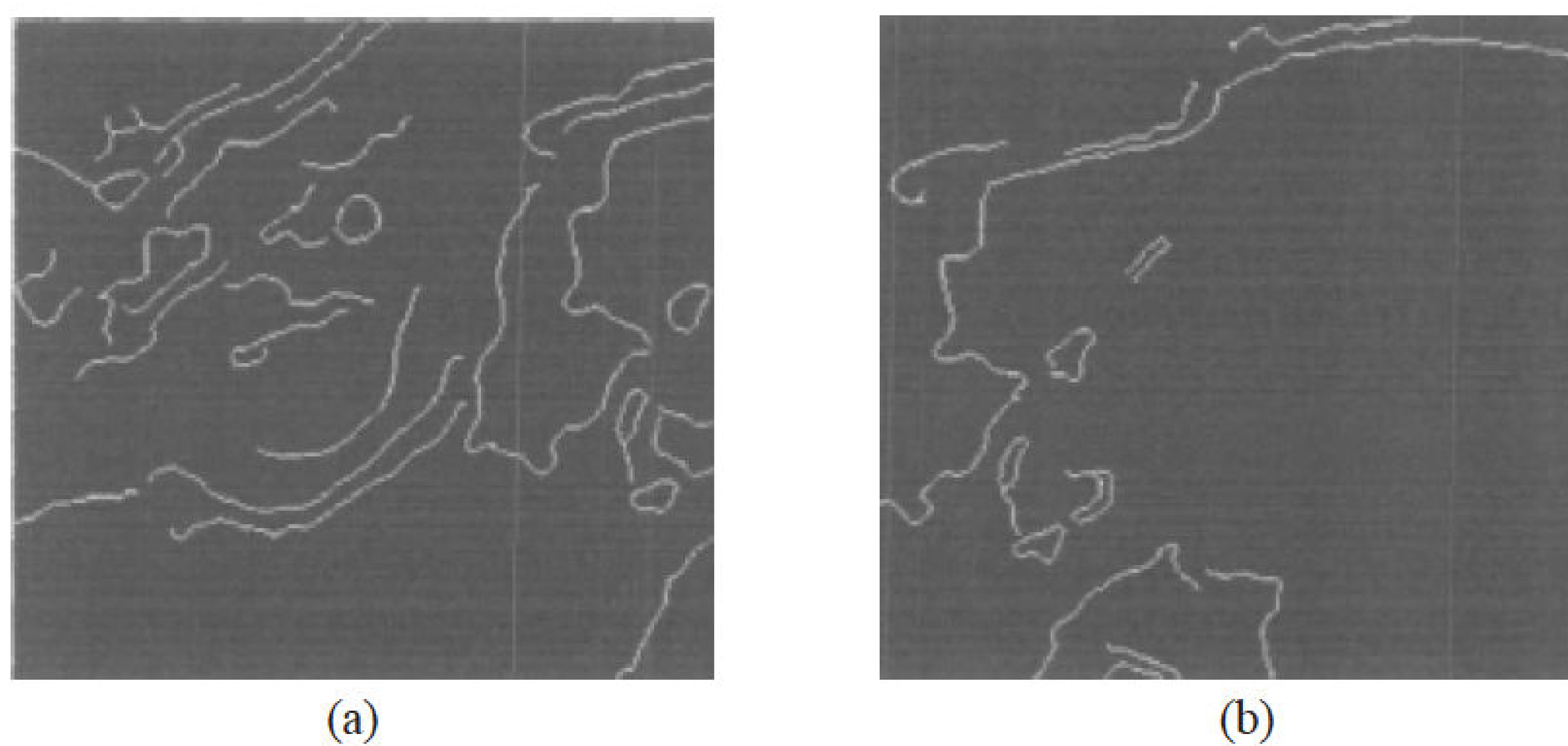
M_1、M_2、M_3。比较M_1、M_2、M_3的大小，取最大点记为M_{max}，则范本中M_{max}点位为原图中p_0点的最佳匹配位置。

原始图像、边缘图像、图像的重叠区域、拼接后图像，分别如图3-8、图3-9、图3-10、图3-11所示。



(a)　　(b)

图3-8　原始图像(250×250)



(a)　　(b)

图3-9　边缘图像

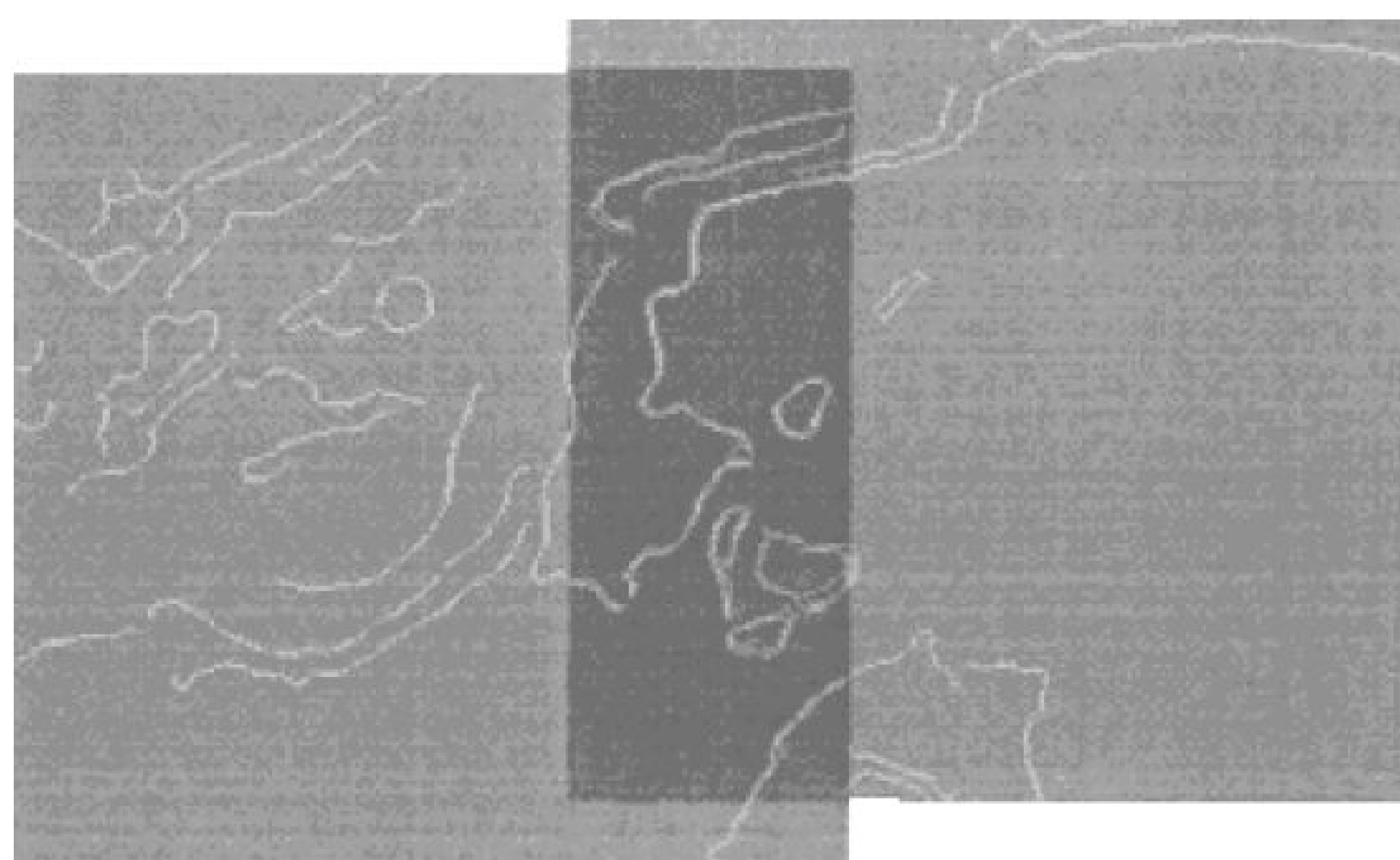


图3-10　图像的重叠区域

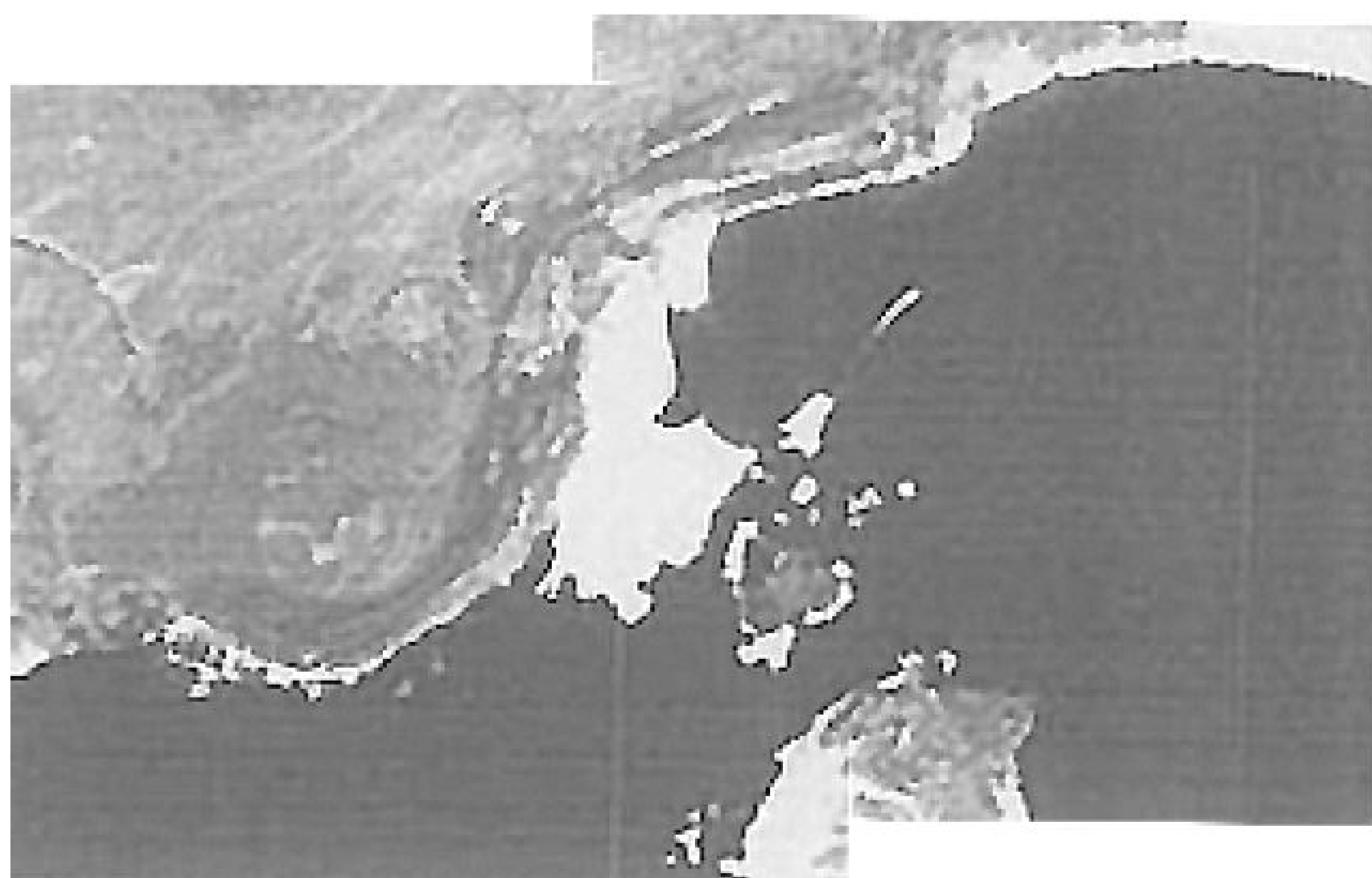


图3-11　拼接后图像

3.4　图像融合

图像融合的目的在于修改拼接图像中过渡区的图像内容，使得两个相邻图像能够平顺地接合在一起，并减少缝合线附近图像的明显程度。图像匹配只依据图像内容找出合适的缝合位置，并不改变图像的内容信息；而图像融合，则依据两个图像本身的内容信息，通过图像融合的操作而产生新的图像内容，此图像内容将具有两个相邻图像在过渡区重叠部分的图像特征，却又能让两个相邻图像看起来自然地接合成一张拼接图像。在图像混合操作时，要在相邻图像的缝合线附近提高图像转变的平滑程度，并且还要让过渡区同时保留两个图像在重叠部分的图像特征。我们分别依照这两种原则，找出最合适的图像融合方式，使得拼接图像具有更好的呈现效果。

为此，一方面以二阶B曲线(2nd-order B-spline)近似为原本过渡区中的灰度变化曲线，使其在微观灰度值的变化细节上产生平顺的变化；另一方面，要同时控制图像的失真与平滑程度，因此利用Snake的能量观念，将曲线本身的平滑、过渡区的失真度和过渡区的平滑转变，以能量的变化为参考，当此参考的总能量越低，越接近过渡区中我们理想的灰度值变化状态。故将此能量以Radientdecent的方法，经过每一次递减，曲线状态逐渐朝总能量降低的方向移动，最后将达到理想的最低状态。

颜色变化的过程要平滑，也就是灰度值的变化程度要趋于平缓，如果灰度值变化的速

度太快，将使颜色的变化中有比较明显的灰度值边缘感觉，导致颜色的转换过程不平缓。颜色转换与灰度值变化，如图3-12所示。

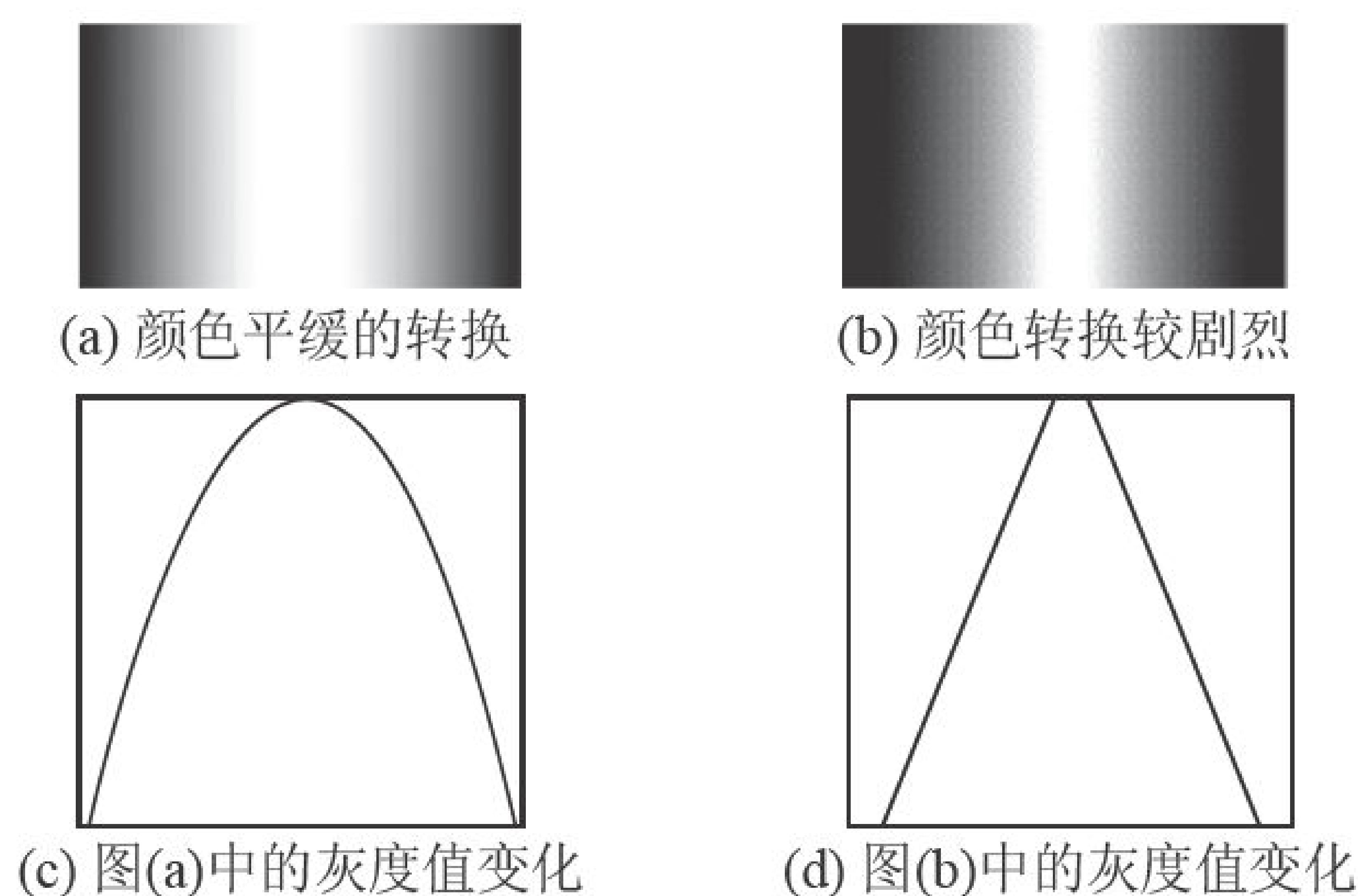


图3-12 颜色转换与灰度值变化

图3-12中(a)与(b)都是在同宽度的范围内由黑到白的灰度值变化，其中图(a)是以图(c)的二次曲线以内插的方式将灰度值由0到255所计算出来的灰度值变化。而图(b)是以图(d)的一次线性所计算出来的灰度值变化，比较图(a)和图(c)，可以发现由二次曲线内插法计算出来的灰度值变化比较平顺，在纯黑色与白色附近仍有平缓的转换，比较图(b)的灰度值变化，在纯黑与白色的附近变化较为剧烈，使得在黑色与白色区域附近有较明显的灰度值边缘感觉。因此，对过渡区中的图像颜色变换希望能以曲线的平滑方式来改变，应该用二阶B曲线本身所具有的特性，使相邻图像的变化更加自然。二阶B曲线是利用基底函数，以控制点的相关坐标内插计算出在控制点范围内的曲线。二阶B曲线的特性在于能够以少数的控制点为基准，通过基底计算出平滑的曲线段，而且经过一连串的控制点，将可以产生一段由许多曲线段所组成的多次曲线，并可以确保曲线段间仍具有平滑变化与连续性。常用的二阶B曲线是以三次曲线来近似为与图像中的某一边缘，以一连串的控制点，其中每4个控制点将产生一个曲线段，相邻的曲线段将共享三个控制点，由此可以产生平滑曲线，希望能精细控制过渡区中灰度值变化的连续曲线，此曲线之间的变化与曲度都将随着内容的不同而有不同复杂度的变化，因此一般三次曲线不容易控制曲线复杂的弯曲。利用二阶B曲线，能对曲线段进行精细的控制，且具有曲线段中的特殊位置与控制点相互转换的便利性，能够更精确地近似为过渡区中的灰度值曲线，使融合后的图像能够调整到理想的效果，我们将利用二阶B曲线找到最近似于过渡区中的灰度值变化曲线。二阶B曲线对于线段的控制，如图3-13所示 。

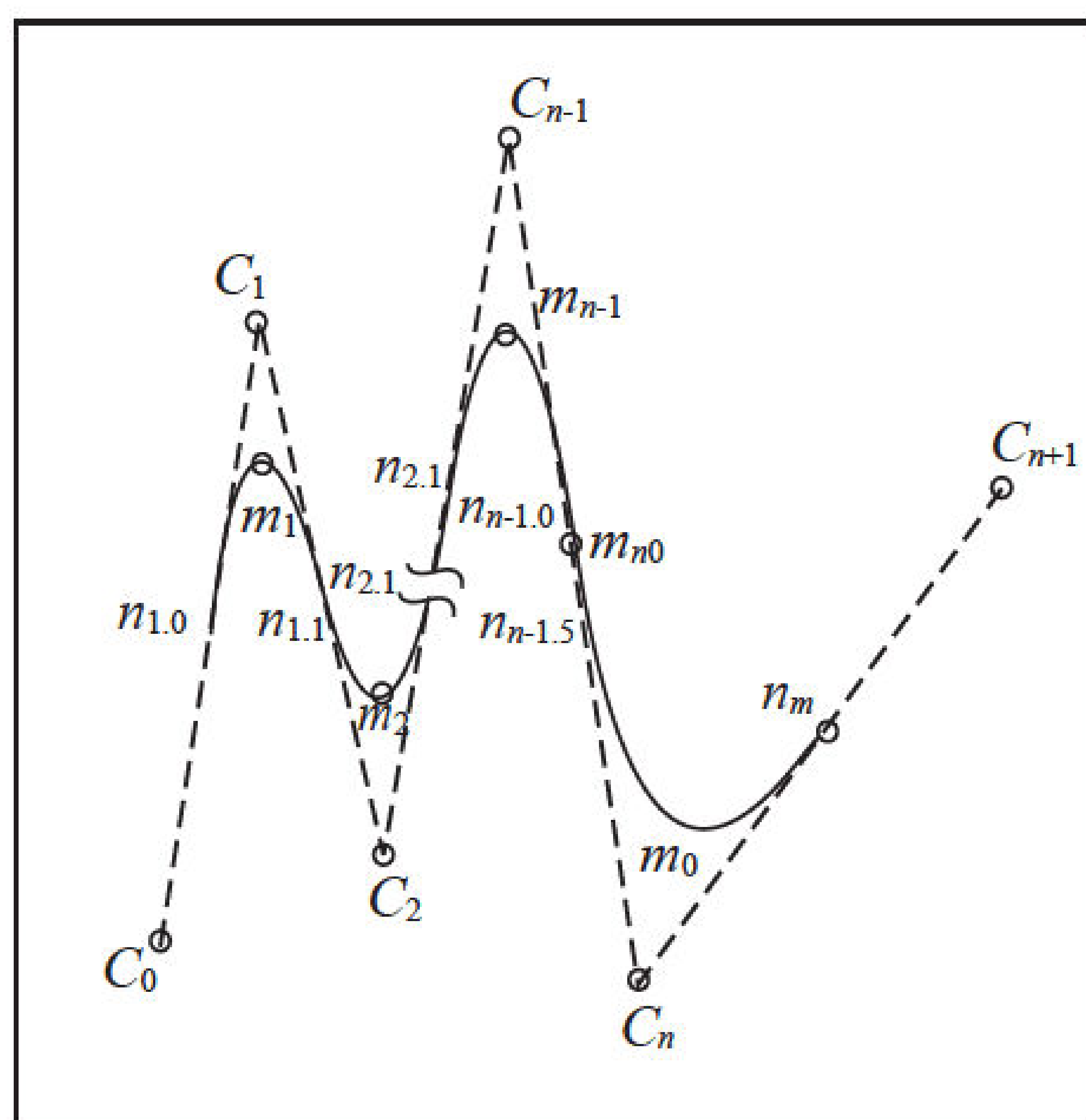


图3-13　二阶B曲线

二阶B曲线，由n+2个控制点，$\{c_j=[x_j, y_j]^T, j=0, 1, \cdots n+1\}$，产生$n$个连接的曲线段，$\{q_j(s)=[u_j(s), v_j(s)]T, j=0, 1, \cdots n\}$。其中每一个曲线段由线性对应的内插参数$s$所得，其值由0到1。二阶B曲线的公式为

$$q_j(s)=[s^2 \quad s \quad 1]\begin{bmatrix}0.5 & -1 & 0.5\\ -1 & 1 & 0\\ 0.5 & 0.5 & 0\end{bmatrix}\begin{bmatrix}c_{j-1}\\ c_j\\ c_{j+1}\end{bmatrix} \tag{3-24}$$

其中$q_j(s)$代表第j段曲线，曲线段是由$c_{j□1}$、c_j、c_{j+1}三个控制点所产生的。依此类推，由一系列的控制点所产生的连续曲线段将可以代表该过渡区中此列灰度值的变化的近似平滑曲线，如公式(3-25)所示，$Q(s)$表示过渡区中以n段曲线所组成的一列灰度值变化曲线。

$$Q(s)=\sum_{j=1}^{n}\sum_{s=0}^{l}q_j(s) \tag{3-25}$$

当s为0时，表示此段曲线j的起点$n_{j,0}$，代入公式(3-24)计算后，曲线段的起点即为前两个控制点$c_{j□1}$和c_j的中点，如公式(3-26)所示。

$$n_{j,0}=q_j(0)=\frac{1}{2}(c_{j-1}+c_j) \tag{3-26}$$

当s为1时，表示此段曲线j的终点$n_{j,1}$，代入公式(3-24)计算后，曲线段的终点即为后两个控制点c_j和c_{j+1}的中点，如公式(3-27)所示。

$$n_{j,1}=q_j(1)=\frac{1}{2}(c_j+c_{j+1}) \tag{3-27}$$

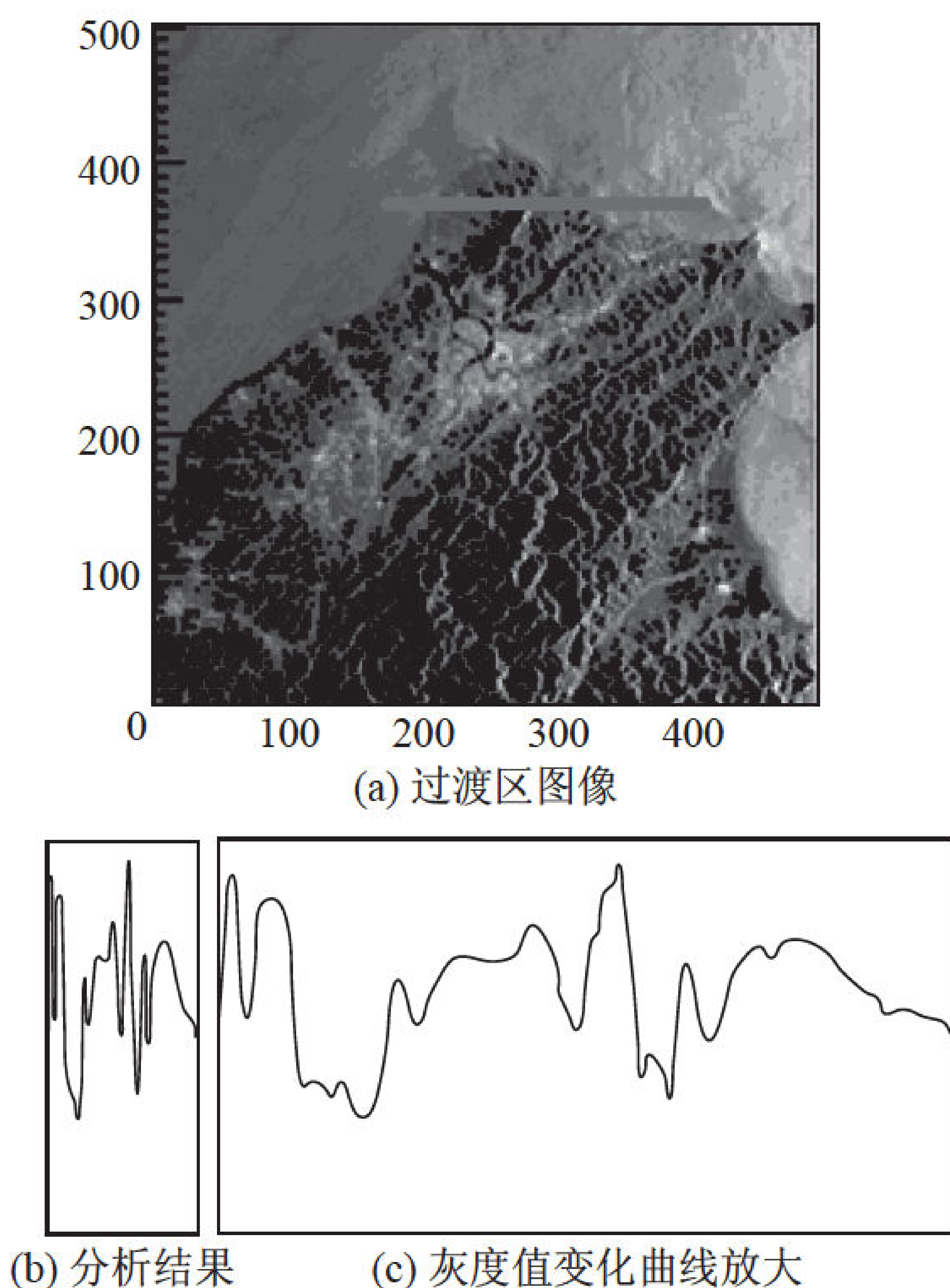
当s为0.5时，表示此段曲线j中曲率最大的地方m_j，将其代入公式(3-24)后，曲线段k中曲率最大的坐标位置为公式(3-28)。

$$m_j = q_j(0.5) = \frac{1}{8}c_{j-1} + \frac{3}{4}c_j + \frac{1}{8}c_{j+1} \tag{3-28}$$

通过式(3-26)、式(3-27)和式(3-28)，可以快速找出曲线段中的特殊位置，而通过曲线段中这三个特征点可以用公式(3-29)逆向找出对应此曲线段的三个控制点的位置。

$$\begin{bmatrix} -2 & \frac{5}{2} & \frac{1}{2} \\ 2 & -\frac{1}{2} & -\frac{1}{2} \\ -2 & \frac{1}{2} & \frac{5}{2} \end{bmatrix} \begin{bmatrix} m_j \\ n_{j,0} \\ n_{j,1} \end{bmatrix} = \begin{bmatrix} c_{j-1} \\ c_j \\ c_{j+1} \end{bmatrix} \tag{3-29}$$

将过渡区的灰度值变化曲线以二阶B曲线来代替，首先要分析出过渡区中相邻图像每列的像素值的变化，以对应过渡区横向像素索引为X坐标，对应索引像素的灰度值为Y坐标。由此可以描绘出此列中的灰度值变化，如图3-14所示。



(a) 过渡区图像

(b) 分析结果　　(c) 灰度值变化曲线放大

图3-14　过渡区灰度值变化图

图3-14中(a)为过渡区中的图像，针对横线部分，即过渡区中第116列的灰度值做分析，其结果为图(b)X方向对应过渡区图像的X坐标，而Y坐标则为灰度值坐标，由0到255。为了方便观察灰度值的变化，我们将图(b)的X坐标放大5倍后变为图(c)。要将二阶B曲线近似为灰度值曲线，由曲线的特性可知，只要是一系列控制点，便可以产生连续且平滑的曲线；此外由于二阶B曲线的特性，曲线段的端点会落在相邻控制点的中点上，所以每个控制点可将其所计算出来的线段控制在两个控制点的边界上。因此，我们只需要在过渡区的灰度值曲线中找到适当的位置作为控制点，则二阶B曲线将能表现连续、平滑且精准的曲线。适当的控制点所在的位置，一是位于灰度值变化的尖端上；二是位于灰度值具有大幅度变化的转角中。针对第一种状况，我们寻找灰度值变化的尖点。当尖点为高峰时，表示与邻近前后的像素值的差距皆为正；当尖点为山谷时，表示与邻近前后像素值的差距皆为负。依此规则，我们可以找出当目前所分析的灰度值坐标$I(x, y)$与前后相邻像素的灰度值的乘积为正时，表示此点为灰度值的尖点。针对第二种状况，当相邻前后像素的灰度值差距达到a时，则此点也应该被取样为控制点。因此，在过渡区的灰度值曲线中选取控制点$c_j(x, y)$的位置为以下三种情况

$$\begin{aligned}&[c_j(x, y)-I(x-1, y)][c_j(x, y)-I(x+1, y)]>0\\&\left|c_j(x, y)-I(x-1, y)\right|>a\\&\left|c_j(x, y)-I(x+1, y)\right|>a\end{aligned} \tag{3-30}$$

灰度值变化曲线的控制点，如图3-15所示。

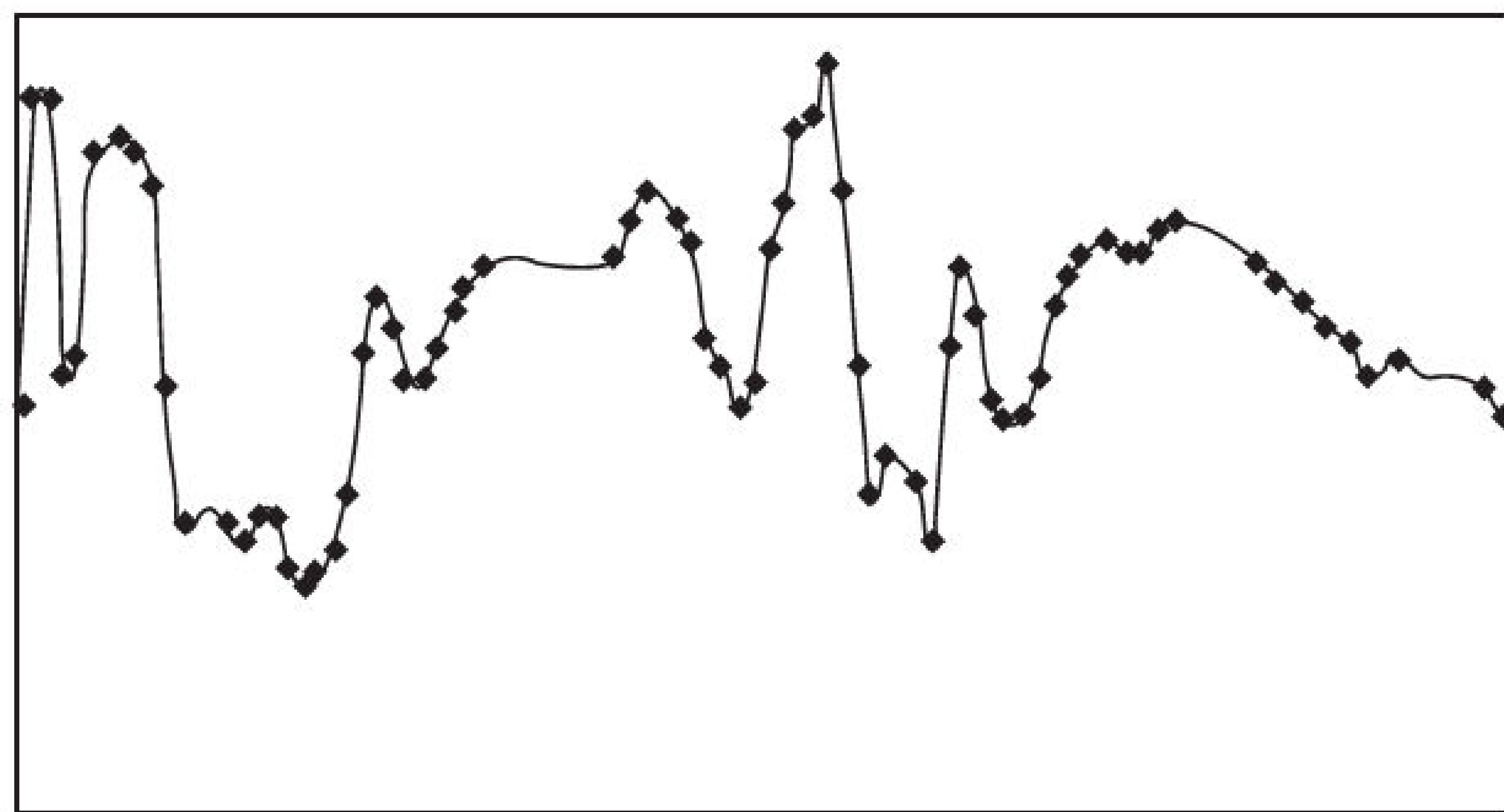
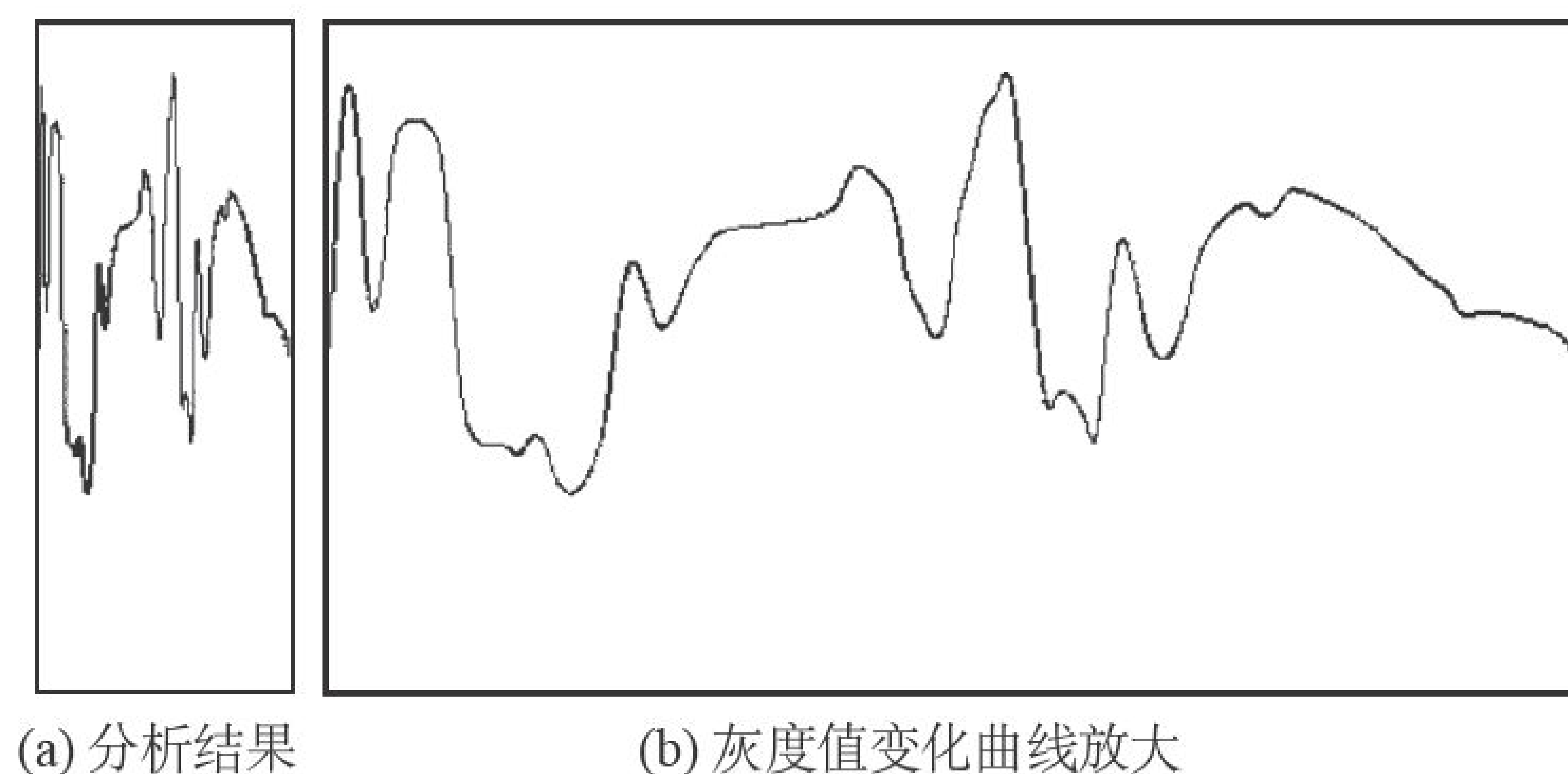


图3-15　灰度值变化曲线的控制点

按图3-14(b)的灰度值变化曲线寻找适当的控制点，其中当a为5时，找到的控制点如图3-15所示，其中所有属于灰度值尖端的部分都被选取为控制点，当以此控制点为中心所产

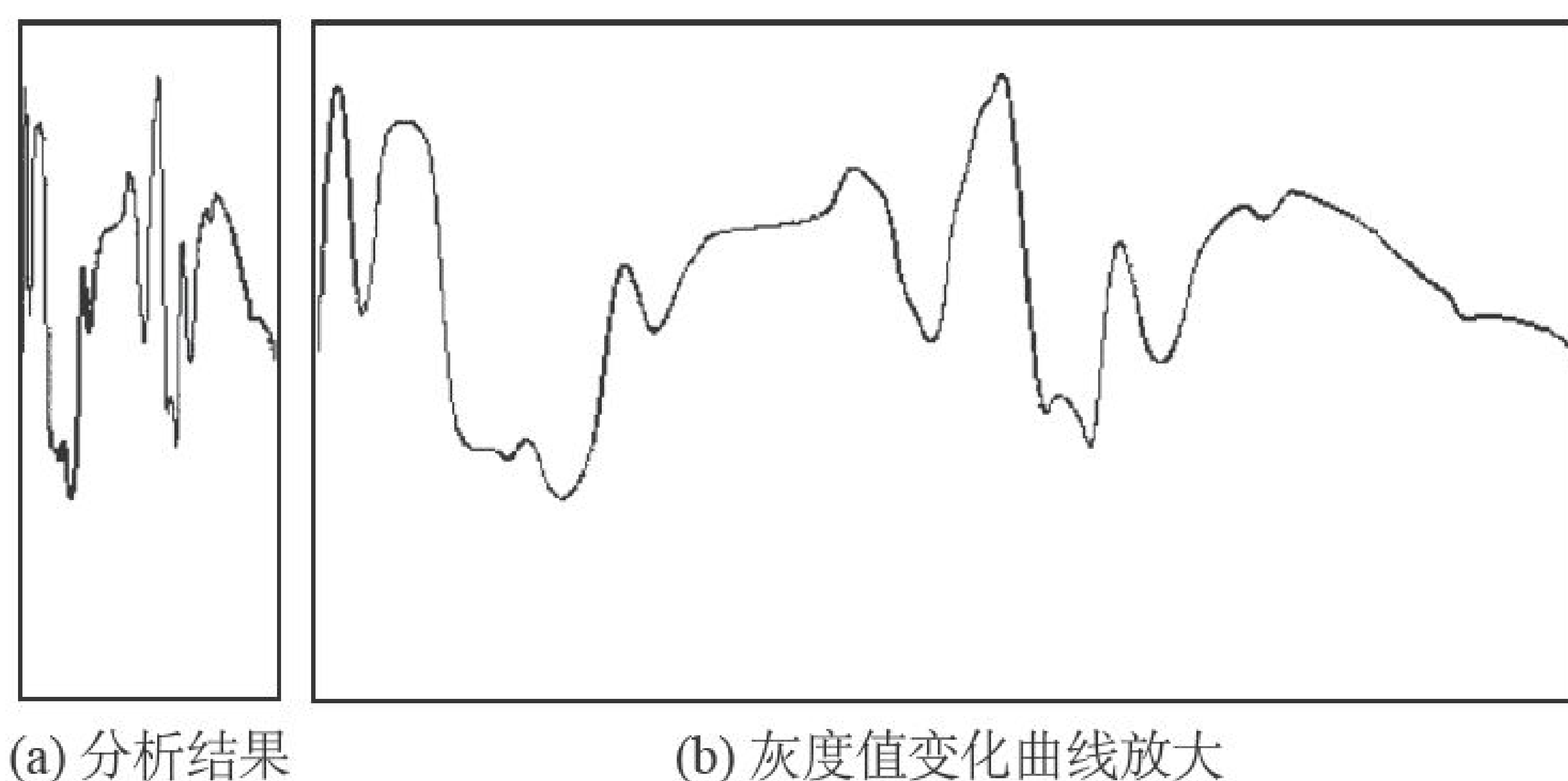
生的对应的二阶B曲线将会使得此尖端部分，包括高峰与山谷，都以内插方式而找到平滑的变化曲线。而相邻灰度值差距超过a，即灰度值相差超过5以上的像素位置，也会被选取作为控制点，而此控制点所对应的二阶B曲线会找到平滑的内插灰度值变化，使得差距较大的灰度值曲线段也能得到平顺的过渡，而避免在过渡区中发生较为剧烈的灰度值变化。由图3-15所找到的控制点分布，利用二阶B曲线的公式(3-30)所产生的灰度值曲线，如图3-16(b)将X坐标放大5倍，可以明显看出，灰度值曲线比图3-14(c)的曲线平滑，但仍保留原灰度值变化的特征，即将会产生较平滑的图像过渡。而图3-16(a)的实际灰度值曲线与图3-14(b)相比较，可以发现一些不平顺的灰度值曲线已经被二阶B曲线的平滑特性所消除，但是仍保留了灰度值曲线的连续变化。



(a) 分析结果　　　　(b) 灰度值变化曲线放大

图3-16　过渡区灰度值变化以二阶B曲线代表，a=5

图3-17中将灰度值差距a调高为10，也就是当相邻像素的灰度值差距达到10，才会被选取作为控制点。由图3-17(b)与图3-14(b)相比较，发现许多灰度值曲线的特征都因为a较大，而被二阶B曲线的平滑特性所消除，失去了灰度值曲线的显著的变化特征，也容易造成过渡区中的图像上下列因为二阶B曲线的平滑程度不一致而产生横向的杂纹。因此，灰度值差距将依靠经验法则，根据过渡区中的内容图像，选择合适的灰度值差距，使其可以保留所需要的灰度值变化特征，又使得灰度值的变化具有平滑的效果。将灰度值变化曲线以二阶B曲线代替，我们将可以利用曲线段中的特殊位置，当s=0.5时曲率最大的m_j来调整图像中的尖点像素值位置，而整个曲线段$q_j(s)$，仍然会因为二阶B曲线的特性而保持平滑，并且还可以由新的位置m_j，反推计算出新的相对应的三个控制点位置，新的控制点位置又会自动取得与邻近曲线段的连续与平滑。因此，我们可以精确地控制过渡区中的灰度值变化程度，又可以确保其图像的平滑程度。



(a) 分析结果　　　　(b) 灰度值变化曲线放大

图3-17　过渡区灰度值变化以二阶B曲线代表，a=10

3.4.1　失真程度

找到一个衡量两个相邻图像在过渡区中的失真度的方法。经由上一节的叙述，将过渡区中的图像灰度值的变动以二阶B曲线来表示，每一曲线段$q_j(s)$都代表了灰度值的变化，因此可以衡量目前过渡区的图像分别与两个图像间在重叠位置上的差别。衡量过渡区中合成的图像与左方、与右方的原始图像的差别程度，同时考虑到图像要能平顺地转变，使得左方的图像在过渡区中越接近左方，需要保留的信息越高，即失真度越低，且在过渡区的最右边则可容忍的失真程度最高。对于右边的图像越靠近过渡区右方则失真度越低，越靠近左边则可容忍的失真度越高。将过渡区中灰度的变化以二阶B曲线近似，分别与左方$L_j(s)$、与右方$R_j(s)$图像在重叠的位置坐标上，以此Blending Function来衡量个别的失真程度，如公式(3-31)所示。

$$\begin{aligned} E_{\text{left},\ j}(s) &= [w_j^l(s) * q_j(s) - w_j^l(s) * L_j(s)]^2, \\ E_{\text{right},\ j}(s) &= [w_j^r(s) * q_j(s) - w_j^r(s) * R_j(s)]^2 \end{aligned} \tag{3-31}$$

其中$w_j^l(s)$与$w_j^r(s)$为对应目前j线段所在s的Blending Function比重分配。$E_{left,\ j}(s)$表示在此j线段灰度值，$q_j(s)$与左方原始图像$L_j(s)$的差距。由Blending Function可以看出，在越靠近过渡区左方的图像所要求的失真程度越低，而$E_{left,\ j}(s)$所计算出来的能量越低，表示过渡区中的图像与左方原始图像的差距越小。同样，对于衡量过渡区图像与右方原始图像的能量$E_{rightt,\ j}(s)$越低，越符合我们对于失真程度的要求。因此，对镶嵌图像在过渡区中的失真度的总能量公式为

$$E_{\text{image},\ j}(s)=E_{\text{left},\ j}(s)+E_{\text{right},\ j}(s) \tag{3-32}$$

过渡区中失真度以能量的方式来衡量，一方面可以评量失真度，另一方面通过Blending Function的控制，使得两个相邻图像在过渡区中具有Blending Function的效果，使得图像在过渡区中能逐渐依照Blending Function平滑地变动而平顺地转变。

3.4.2　能量总和

为了能够使镶嵌图像过渡区中的灰度值的曲线接近我们理想中能够呈现图像的平滑转换和降低两个图像各自的失真度，我们采用图像分割中Active Contour Model的方式，以能量观念来调整灰度值曲线的位置，最后在达到最小能量时，使得灰度值曲线达到最理想的变化。其中以Snake[36][37]为最典型的方法。Snake最早应用于图像分隔，要在图像中找到最接近图像内容的轮廓，应先给定封闭曲线的初始位置，接着将此曲线以内部能量(Internal Energy)和外部能量(External Energy)来决定Snake曲线的移动方向与位置。Snake将曲线订定为$v(s)=[x(s),\ y(s)]$且s为订定此曲线的参数，其值由0到1。该曲线的总能量公式为

$$E_{\text{snake}}=\int_0^1 E_{\text{int}}[v(s)+E_{\text{ext}}v(s)]\mathrm{d}s \tag{3-33}$$

内部能量来自封闭曲线本身的平滑程度与连续的特性，其能量计算公式为

$$E_{\text{int}}=[a(s)\left|v_s(s)\right|^2+\beta(s)\left|v_{ss}(s)\right|^2]/2 \tag{3-34}$$

公式(3-34)中$v_s(s)$为此曲线的一次微分，用来表示此曲线的连续程度；而$v_{ss}(s)$为此曲线的二次微分，用来表示此曲线的弯曲程度。因此，内部能量表示此曲线本身轮廓的样子。当该能量越低，则曲线将表现得连续且平顺少弯曲。

外部能量则是根据外在图像信息，而影响曲线段变化的能量。一般常用的外部能量公式为

$$E_{\text{ext}}=-\left|\nabla I(x,\ y)\right|^2 \tag{3-35}$$

其中$I(x,\ y)$是指图像灰度梯度值，即图像的灰度值变化的程度。

一般外部能量用于判断图像中的边缘位置。图像在边缘位置的灰度梯度值将呈现较大的变化，越明显的边缘则灰度梯度值越大。外部能量越小，则表示目前曲线所在的位置处于越明显的图像边缘位置。Snake是利用这两种能量，使得曲线维持平滑与连续的状态下，能向图像中的边缘靠近。Snake在能量最小化的过程中，达到能量稳定且最小的状态

时，表示此曲线已经将图像中的内容以曲线圈选出来。在此最小化的过程中，曲线不断地移动，向理想的轮廓趋近，其过程如同蛇一样渐渐扭动逼近最佳的图像区域，因此命名为Snake。我们一方面希望过渡区中的灰度值变化以曲线的样子来分析时，曲线本身可以越接近平滑使得合成的图像更自然，也使得相邻图像间的过渡较为平顺；另一方面希望依照失真程度能够按Blending Function的比重分配来控制此曲线，使曲线能处于满足Blending Function的比例分配，又能够处在相对于两个图像失真度最小的位置。因此，我们依照Snake的能量公式来计算过渡区中的灰度值曲线而通过能量的递减，逐渐接近理想的图像呈现效果。

图像的平滑程度，一方面来自控制失真程度的Blending Function，另一方面来自二阶B曲线本身的平滑程度。曲线越平滑，此曲线所呈现的灰度值转换越平顺；曲线的变化越激烈，则此灰度值的变动程度会随之呈现剧烈的变化，而使图像呈现较锐利的效果。Snake中的内部能量，主要是控制曲线的连续性与弯曲程度，因此以二阶B曲线的连续及平滑的特性来取代其内部能量，而形成所谓的B-Snake[44][45]，因而能量最小化时，可以只考虑其外部能量的图像，内部能量则由二阶B曲线的特性来维持曲线的平滑。然而，我们所要考虑的不仅是曲线的连续性，还有曲线本身的平滑程度，也就是相邻灰度值的变化程度。因此，我们对于曲线本身在图像合理的转换及减小失真度的条件下，越平滑越好。而曲线的平滑程度可用曲线一阶微分，即曲线的斜率来判断。斜率越小，表示曲线的变化越平缓。我们以累积二阶B曲线中的斜率为Snake的内部能量，如公式(3-36)所示。

$$E_{\text{int}}(s)=\sum_{s=0}^{1}\left|q_j{}'(s)\right| \tag{3-36}$$

在公式(3-36)中，由于每段曲线$q'_j(s)$中的变量由参数S所控制，因此要测量曲线$q'_j(s)$的平滑程度，就必须对S做微分，由此我们可以得到公式(3-37)。

$$q_j{}'(s)=2(\frac{1}{2}*c_{j-1}-c_j+\frac{1}{2}*c_{j+1})*s+(c_j-c_{j-1}) \tag{3-37}$$

而对于Snake的外部能量，我们选取图像的失真度来衡量。失真度是由二阶B曲线所代表的灰度值曲线与原始的两个图像间的灰度值差异所组成的。因此，我们以公式(3-34)衡量图像的失真程度，累积曲线段中对两个相邻图像在过渡区中的失真程度，为所要依据的外部能量，如公式(3-38)所示。

$$E_{\text{ext}}(s)=\sum_{s=0}^{1}[w_j^l(s)*q_j(s)-w_j^l(s)*L_j(s)]^2+[w_j^r(s)*q_j(s)-w_j^r(s)*L_j(s)]^2 \tag{3-38}$$

衡量过渡区中图像混合的能量公式为

$$E_{\text{snake}} = \sum_{j=1}^{n}\sum_{s=0}^{1} \lambda * E_{\text{int}}[q_j(s)] + E_{\text{ext}}[q_j(s)] \tag{3-39}$$

其中λ表示内部能量与外部能量的参数比重，因为内部能量过大或过小都将影响图像的失真程度与转换的平滑度，所以λ能够决定两种能量所占的比重。能量越小，表示由二阶B曲线所形成的灰度值变化将会越平缓，且过渡区中的图像与两个相邻图像的失真度最小。

3.4.3　能量最小化

首先以二阶B曲线近似于过渡区相邻图像重叠位置的灰度值之平均所形成的灰度值曲线(见图3-18)为初始位置而计算能量递减。

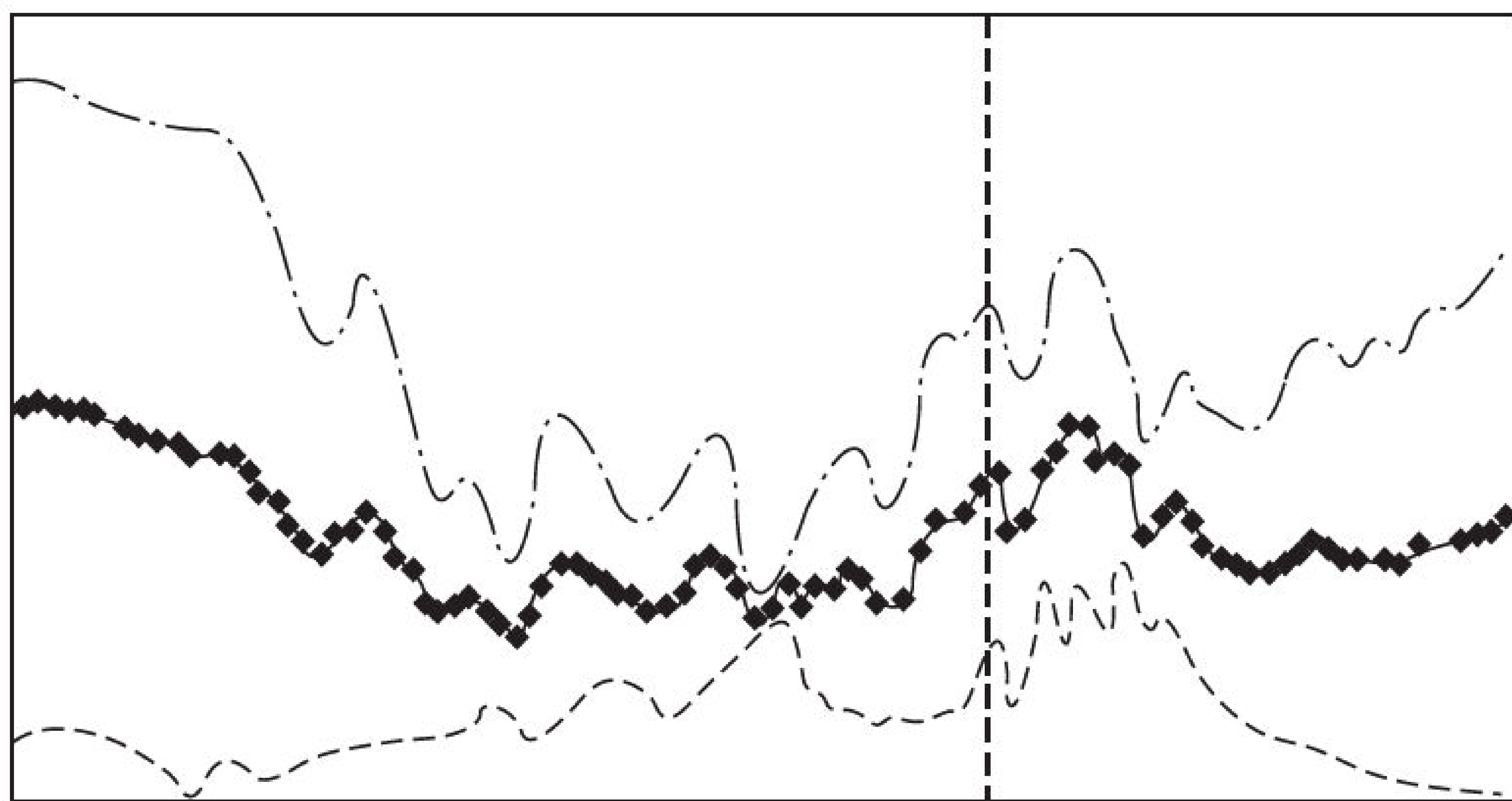


图3-18　二阶B曲线初始位置

在上节中找到的能量总和越小，表示越接近理想中的灰度值混合效果，因此我们利用逐渐递减(Gradient Decent)的数值方法来找到公式(3-39)的能量最小化，也就是找到此能量函式的递减方向，将二阶B曲线每次都趋向能量递减的方向，最后到达会使能量最小的曲线位置，即能量将稳定最小化，如公式(3-40)所示。

$$q_j(s)' = q_j(s) - \frac{\partial E_{\text{snake}}}{\partial q_j(s)} \tag{3-40}$$

可以分别对外部能量与内部能量做偏微分，如公式(3-41)所示。

$$\frac{\partial E_{\text{snake}}}{\partial q_j(s)} = \frac{\partial E_{\text{int}}}{\partial q_i(s)} + \frac{\partial E_{\text{ext}}}{\partial q_i(s)} \tag{3-41}$$

找到每次二阶B曲线所应该递减的量值，直到下次产生新的$q_j(s)'$与当前的$q_j(s)$相同或到设定的最小差距时，即找到理想的灰度值混合状态。为了便于表示，将公式(3-24)的线段公式展开为

$$q_j(s)=(\frac{1}{2}*c_{j-1}-c_j+\frac{1}{2}c_{j+1})*s^2+(c_j-c_{j-1})*s+\frac{1}{2}(c_{j-1}+c_j) \tag{3-42}$$

其中，$A_j=(\frac{1}{2}*c_{j-1}-c_j+\frac{1}{2}*c_{j+1})$；$B_j=(c_j-c_{j-1})$；$C_j=\frac{1}{2}*(c_{j-1}+c_j)$。

要将公式(3-42)对$qj(s)$做偏微分，我们可以将内部能量与外部能量分别对$qj(s)$做偏微分，之后再将其结果结合起来。将内部能量对$qj(s)$做偏微分，得到公式(3-43)。

$$\frac{\partial E_{\text{int}}}{\partial q_j(s)}=\lambda' A_j \tag{3-43}$$

其中$\lambda'=2\lambda$。

公式(3-43)表示二阶B曲线要趋于平滑每次递减会依照当前的三个控制点坐标而取得适当的量值，直到曲线达到水平线，则不再变动。外部能量中，由前半段衡量左半图像的能量对$q_j(s)$做偏微分得到公式

$$\begin{aligned}&\frac{\partial[w_j^l(s)*q_j(s)-w_j^l(s)*L_j(s)]^2}{\partial q_j(s)}\\&=2[w_j^l(s)]^2[q_j(s)-L_j(s)]\end{aligned} \tag{3-44}$$

对于外部能量中的后半段衡量右半图像的能量对$q_j(s)$做偏微分得到公式

$$\begin{aligned}&\frac{\partial[w_j^r(s)*q_j(s)-w_j^r(s)*R_j(s)]^2}{\partial q_j(s)}\\&=2[1-w_j^r(s)]^2[q_j(s)-R_j(s)]\end{aligned} \tag{3-45}$$

由公式(3-43)、公式(3-44)、公式(3-45)整理之后可以得到最后的能量最小化的结果，即

$$\begin{aligned}&\frac{\partial E_{\text{snake}}}{\partial q_j(s)}=\lambda' A_j+2\{[2(w_j^l(s'))^2-2w_j^l(s')+1]q_j(s)-\\&w_j^l(s')L_j(s)-[1-w_j^l(s')]R_f(s')\}\end{aligned} \tag{3-46}$$

由公式(3-46)可以确定一段二阶B曲线在每一次递回时应该减少的能量值，直到两次递回的差别小于误差值e时，即达到此线段在过渡区中可以呈现理想的灰度值变化。公式

(3-43)、公式(3-44)、公式(3-45)所产生的三个新的曲线段的特殊位置，即可产生对应的三个新控制点位置，使得连接的两个曲线段中共享相邻的两个控制点也会随之更新。因此每递归一次，除了会使当前线段的能量带往稳定的低点，也会影响相连接的两段曲线段。因此只要依序对各个线段做能量最小化的处理，则过渡区中的二阶B曲线也会将整条曲线带往最低能量之处，而不是只有单一的曲线段。

利用此特性，以过渡区为起始点，每段二阶B曲线将依序递回，直到此列中所有的曲线都达到能量最小化，之后再换下一列，直到过渡区所有的像素列都完成二阶B曲线的能量最小化后，则过渡区的图像混合完成。

以图3-12(a)的过渡图像中的一段二阶B曲线为例，按λ'=0.2、误差控制ε=0.1所做的能量递减，如图3-19所示。此线段处于缝合线以左，因此二阶B曲线会因为能量递减而逐渐接近左边图像的灰度值变化。由图3-20中，可以看出其能量误差逐渐减少，经过9次递回达到稳定的状态，即两次递回的相差小于ε。

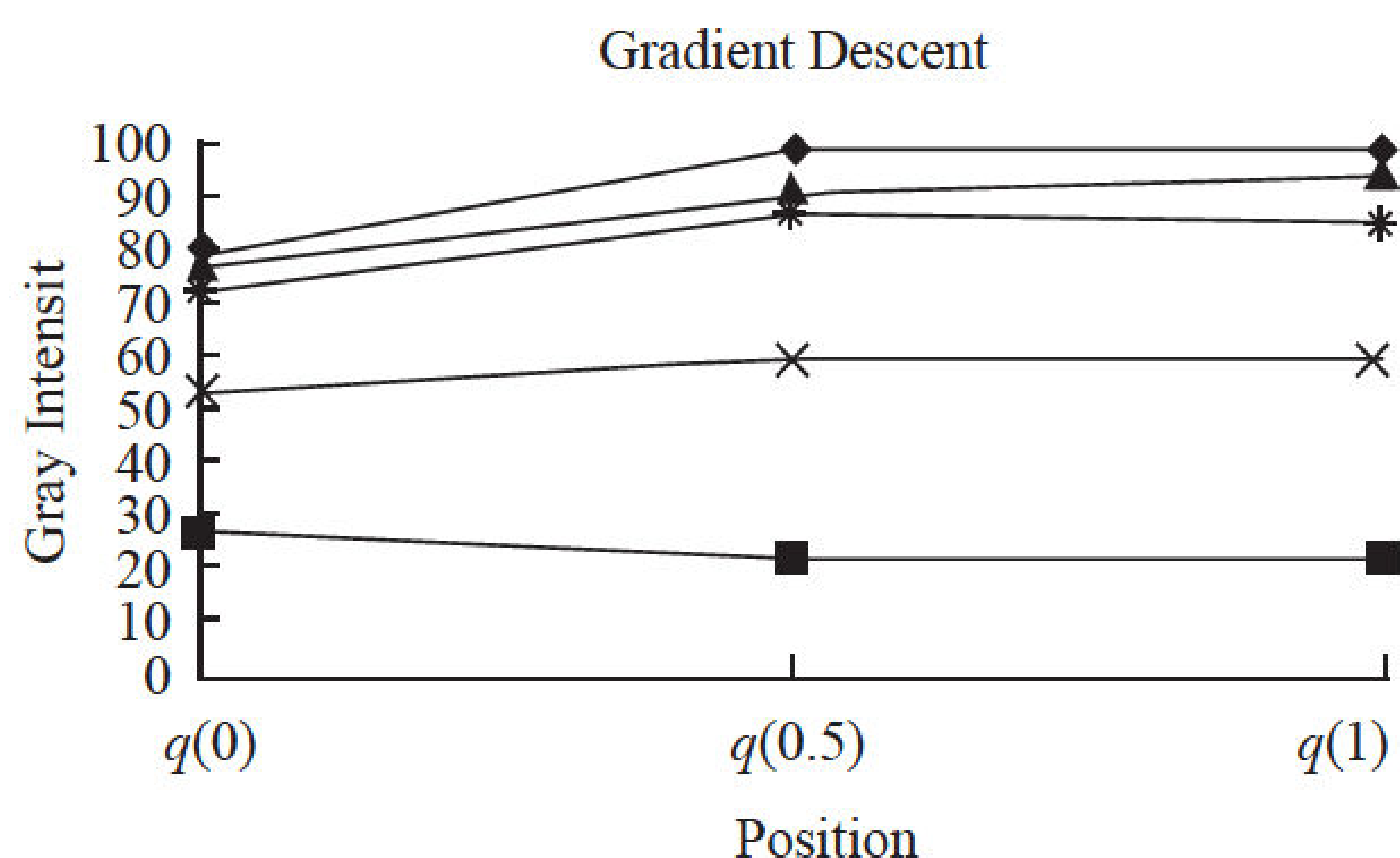


图3-19　曲线渐进

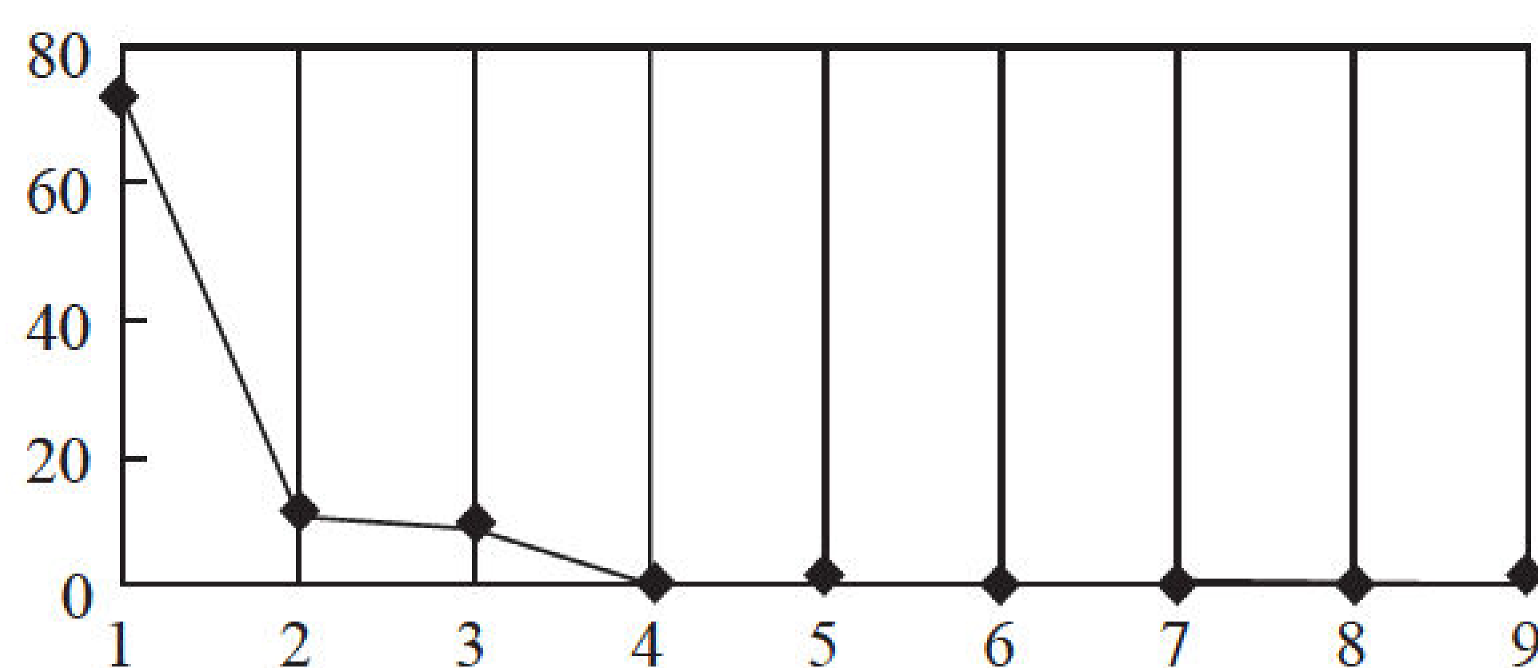


图3-20　递减误差

对于内部能量λ'，我们以0和0.2为参数代入。当λ'为0时，表示内部能量不起作用，即

不会将曲线段的平滑程度当作总能量来考虑；当λ'为0.2时，可由图3-21看出，曲线将渐趋平滑，而依此参数所得到的镶嵌图像也呈现较平滑的效果。

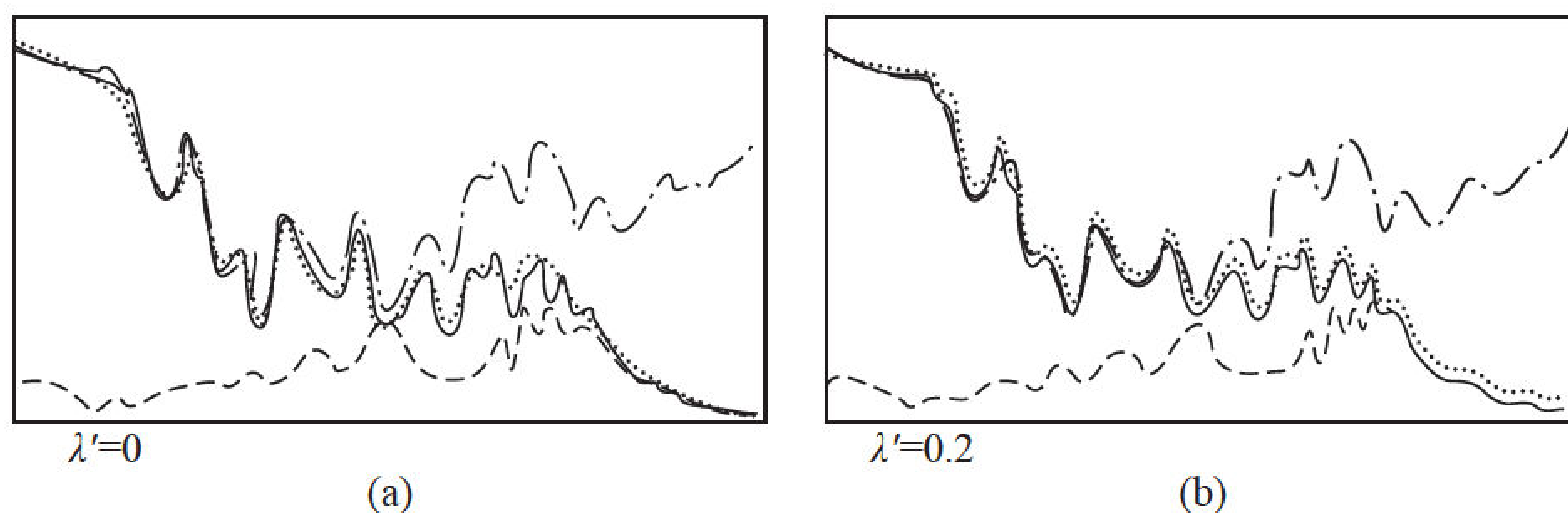


图3-21　内部能量参数比较

利用二阶B曲线的能量最小化来找到最佳的图像混合效果，其中包含控制图像的转换平滑和减小失真度。寻找最佳图像混合，而使用二阶B曲线能量最小化的算法如下所述。

(1) 首先以二阶B曲线近似于过渡区相邻图像重叠位置的灰度值之平均所形成的灰度值曲线为初始位置而计算能量递减。

(2) 从过渡区中的第一列开始，由过渡区的起始点向终点，对灰度值曲线的每一段二阶B曲线按公式(3-26)做能量递减。

(3) 重复步骤(2)，直到过渡区中的图像的灰度值皆以二阶B曲线达到能量最小的稳定状态。

(4) 此时过渡区中的图像就是完成的图像混合。

原始图像I_1、原始图像I_2、拼接后图像、经过平滑后的拼接图像，分别如图3-22、图3-23、图3-24、图3-25所示。



图3-22　原始图像I_1



图3-23　原始图像I_2



图3-24　拼接后图像



图3-25　经过平滑后的拼接图像

3.5　实验结果与分析

选取两幅连续的、由无人机拍摄的原遥感图像，大小为370×370，256级灰度航拍图像，如图3-26、图3-27所示。



图3-26　原图像I_1



图3-27　原图像I_2

相位校正+特征融合的拼接结果、本书所采用方法的拼接结果，分别如图3-28、图

3-29所示。



图3-28　相位校正+特征融合的拼接结果



图3-29　本书所采用方法的拼接结果

连续16幅序列图像及其拼接结果，分别如图3-30、图3-31所示。

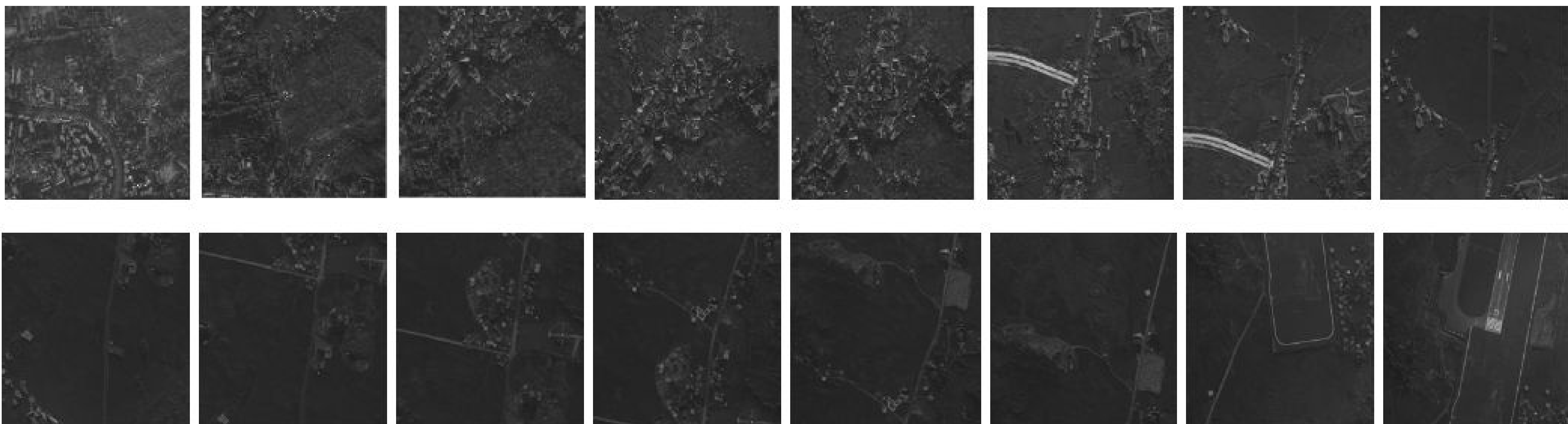


图3-30　连续16幅序列图像

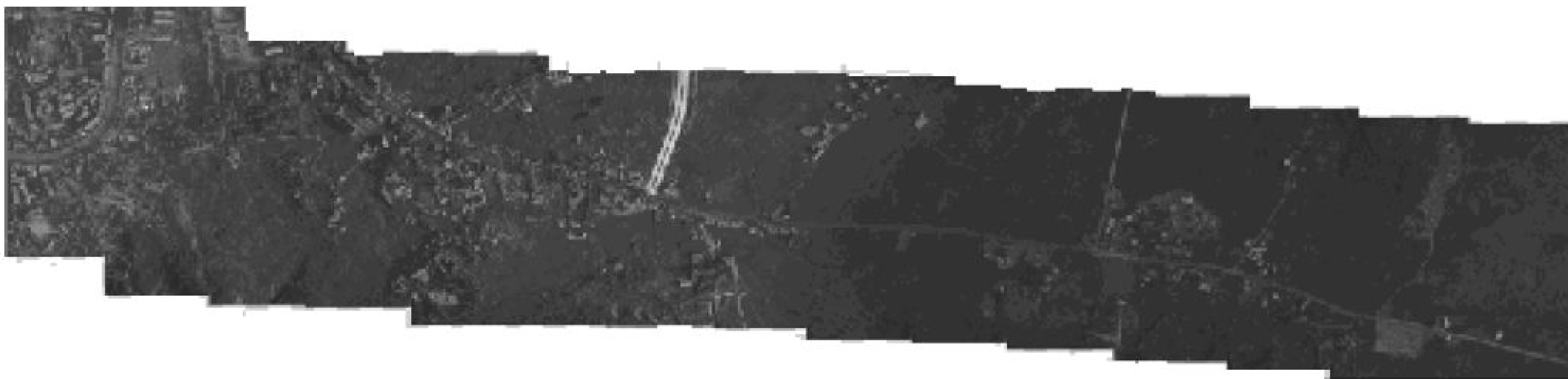


图3-31　连续16幅图像的拼接结果(旋转90度后显示)

各种拼接算法性能的比较，如表3-1所示。

表3-1　各种拼接算法性能比较

项目 \ 算法	Szeliski算法	Imad算法	本书算法
特征	Texture	Harris Corner	Edge
计算时间/s	5.36	5.53	4.68

3.6　结论

本章提出一种小波变换与Canny算法相结合的遥感图像大边缘检测方法，该方法是为图像配准而设计的。该方法将小波变换多分辨率分析特性和Canny算法检测边缘思想相结合，提取遥感图像中大的边缘，确定两幅图像的重叠范围，提取具有丰富图像信息的模板，设计模板匹配搜索策略，完成重叠图像区域的粗匹配和细匹配，得到两幅图像相对的位置偏移。根据图像匹配的计算结果，完成前后两幅图像的对接，并对图像拼接过渡区域的幅度值平滑化。仿真实验研究证明了算法的有效性。两幅图像之间拼接的时间较快，最大连续拼接图像数量为36幅，基本可以满足实际应用的要求。

但是这种拼接技术本身也存在一些问题。如果仔细观察图3-29，就会发现在边缘地方还是存在少量的锯齿，这是因为多分辨率拼接技术存在的不足可能造成局部的瑕疵，这是由计算中的数据近似造成的，因此难以克服。多分辨率拼接存在的另一个不足是合成后的图像会变暗和模糊。在图3-27中本来很清楚的纹理，到了图3-29就会变得模糊。产生这种现象的原因是经过多次高斯滤波处理后信号因近似而减弱，图像也因此变暗和模糊。多分辨率拼接技术没有考虑纹理细节的保持。另外，对于特殊条件下拍摄得到的图像不能很好地实现拼接处理，主要包括以下两方面。

(1) 对于无人机大角度倾斜姿态条件下拍摄得到的图像不能实现正确拼接，原因是在角度倾斜的飞行条件下，相邻两幅图像之间的重叠区域很小，且存在较大的图像畸变，使计算过程中的特征提取及匹配搜索无法顺利进行。

(2) 对于在大面积云层遮挡的情况下拍摄得到的图像不能实现正确拼接。云层的遮挡使不同视角的图像存在景物差异，导致匹配计算不能稳定实现。

这些问题将在以后的研究中加以解决。

第4章

颜色空间与颜色相似性度量



与灰度图像相比，彩色图像由于具有更为丰富的可视化信息，因此得到更为广泛的应用。彩色图像中像素颜色通常由三维颜色空间向量来表示，因此与灰度图像相比，在图像的表达方式和处理方法上存在较大的差异，使得彩色图像拼接具有更大的难度。为了更好地实现彩色图像的拼接处理，本章要分析几种常用且具有代表性的颜色空间，分别给予相对详细的介绍。由于RGB颜色空间是最为常用的一种颜色空间，大多数用于彩色图像的输入和输出。由于RGB颜色空间本身的缺陷，在彩色图像处理中经常把RGB颜色空间转换到其他颜色空间进行处理。颜色空间转换是研究设备彩色特性和进行色彩管理工作的基础。本章对RGB颜色空间与其他各类颜色空间的转换方法进行详细的介绍。另外，HSI颜色空间是相对符合人眼颜色视觉特性的颜色空间，在彩色图像处理中得到广泛的应用，但在颜色空间变换过程中同样存在复杂的计算，影响了彩色图像处理的实时性。根据HSI与RGB颜色空间的转换公式，推导出两种空间各颜色分量之间的一种数值对应关系，据此，定义了色度饱和度相似系数、亮度相似系数。根据人眼在分辨颜色时的特性，最终定义颜色相似系数，确定两种颜色相似的判别规则。通过实验证明了该方法符合人眼颜色相似判别规律，计算简单，易于实现，可用于彩色图像匹配中的像素颜色相似性计算。

4.1 颜色空间分析

4.1.1 面向设备的颜色空间

面向设备的颜色空间适合在输出显示场合下使用。计算机彩色输入/输出设备使用与设备相关的颜色空间，如CRT显示器和扫描仪使用RGB颜色空间，打印机使用CMYK颜色空间，但各种外部设备使用的颜色空间并不一致。因此，除了分析研究它们各自所在的颜色空间外，还必须清楚它们之间的相互转换关系，这在计算机系统彩色管理工作中十分重要。[136]

1. RGB颜色空间

光谱中最重要的颜色是红(R)、绿(G)、蓝(B)三基色，所有颜色都可以由三基色相加而产生。RGB颜色空间是一个立方体三维坐标空间结构，分别用红、绿、蓝表示三个坐标轴，如图4-1所示。立方体的底部$R=G=B=0$处为黑色，顶部与其相对角$R=G=B=255$处为白

色。由于图像采集和显示设备使用的是RGB颜色空间，所以RGB颜色空间是彩色图像处理中最基础、最常用的颜色空间。

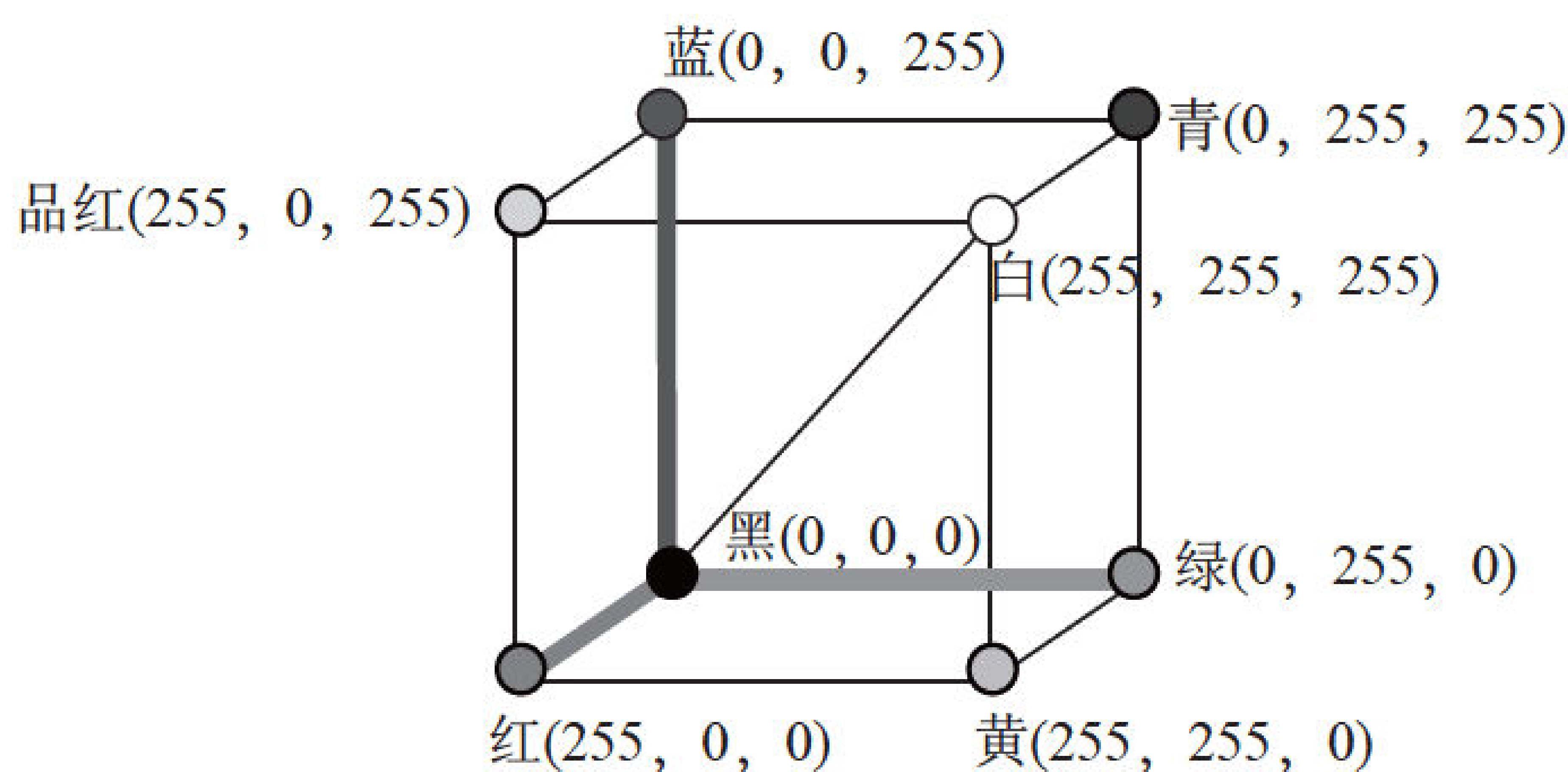


图4-1　RGB颜色空间

RGB颜色空间的主要缺点：①不直观，即从RGB值中很难看出其所表示的颜色的认知属性；②不均匀，即两个色点之间的距离不等于两个颜色之间的知觉差异；③对硬件设备具有依赖性。因此，RGB颜色空间是一个与设备相关的、颜色描述不完全直观的空间。为了克服RGB颜色空间的不均匀和不直观的缺点，在彩色图像处理中大多采用更加符合颜色视觉特性的颜色空间。RGB颜色空间能被转变为所需要的其他任何颜色空间。由于任何三个基色都能表示一种颜色，根据三基色理论的实现要求，可以使用不同颜色空间的不同基色来表达同一种颜色。[147]颜色空间变换提供了一种三基色颜色空间向另一种三基色颜色空间的映射方法，实现从一组原色向另一组原色转换，这是由于任何原色刺激都可以由其他组原色刺激的混合来生成。例如，从RGB到R'G'B'变换可以写为

$$R'=a_{11}R+a_{21}G+a_{31}B \tag{4-1}$$

$$G'=a_{12}R+a_{22}G+a_{32}B \tag{4-2}$$

$$B'=a_{13}R+a_{23}G+a_{33}B \tag{4-3}$$

或者写为

$$P'=TP \tag{4-4}$$

这里T变换矩阵中，P和P'是不同的颜色空间。

$$T=\begin{pmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{pmatrix} \tag{4-5}$$

2. CMY(CMYK)颜色空间

CMY颜色空间是一个矩形直角空间结构，其三基色分别为深蓝(C)、品红(M)和黄(Y)。它利用减色法获得各种颜色，常用于非发射显示，如喷墨式打印机等。等量的颜料三基色可以产生黑色，但实际上这样产生的黑色是不纯的。通常为了产生真正的黑色，要加入第四种颜色——黑色(K)，提出了CMYK颜色空间。CMY颜色空间不直观，而且是非线性感知的。CMY颜色空间的三基色与RGB颜色空间的三基色为互补色。虽然这两个三基色集合产生的空间并不完全真正对应，但有简单的近似公式可以相互变换，从RGB颜色空间到CMY颜色空间的变换公式为

$$\begin{aligned} C &= 1-R \\ M &= 1-G \\ Y &= 1-B \end{aligned} \tag{4-6}$$

从CMY颜色空间到CMYK颜色空间的变换公式为

$$\begin{aligned} K &= \min(C，M，Y) \\ C &= \frac{C-K}{1-K} \\ M &= \frac{M-K}{1-K} \\ Y &= \frac{Y-K}{1-K} \end{aligned} \tag{4-7}$$

3. YUV与YIQ颜色空间

YIQ、YUV和YC_rC_b颜色空间都产生一种亮度分量信号和两种色度分量信号，而每一种变换使用的参数都是为了适应某种类型的显示设备。其中，YIQ适用于NTSC彩色电视制式，YUV适用于PAL和SECAM彩色电视制式，而YC_rC_b适用于计算机的显示器。

在彩色电视制式中，使用YIQ和YUV模型来表示彩色图像。YIQ颜色空间是美国国家电视标准委员会(NTSC)定义的电视广播标准。在YIQ颜色空间中，Y为亮度信息，I、Q为色度值，是两个彩色分量。其中，I为橙色向量，Q为品红色向量，各分量近似正交。正交颜色空间主要应用在彩色电视系统，可以减少电视信号传输的带宽，是彩色图像编码和压缩的基础。欧洲定义了相交替制式(PAL)，使用YUV格式。在这种颜色空间中，Y为亮度信息，U、V为色差信号，U、V是构成彩色的两个分量。YUV颜色空间与YIQ颜色空间类似，差别仅在于空间多了一个33°的旋转，目的是可以与黑白电视兼容。电视信号在发射

时，转换成YUV形式，接收时再还原成RGB三基色信号，由显像管显示。YUV表示法中的亮度信号(Y)和色度信号(U、V)是相互独立的，即Y信号分量构成的黑白灰度图与用U、V信号构成的另外两幅单色图是相互独立的。此外，黑白电视能接收彩色电视信号，就是利用了YUV分量之间的独立性。YUV表示法的另一个优点是可以利用人眼的特性来降低数字彩色图像所需要的存储容量。

1) YUV与RGB的空间变换

在考虑人的视觉系统和阴极射线管(CRT)的非线性特性之后，RGB和YUV的对应关系可以近似地用下面的方程式表示

$$\begin{aligned} Y &= 0.299R + 0.587G + 0.114B \\ U &= -0.147R - 0.289G + 0.436B \\ V &= 0.615R - 0.515G - 0.100B \end{aligned} \tag{4-8}$$

2) YIQ与RGB的空间变换

RGB和YIQ的对应关系用下面的方程式表示

$$\begin{aligned} Y &= 0.299R + 0.587G + 0.114B \\ I &= 0.596R - 0.275G - 0.321B \\ Q &= 0.212R - 0.523G + 0.311B \end{aligned} \tag{4-9}$$

4.1.2　面向视觉感知的颜色空间

面向设备的颜色空间与人眼的视觉感知有一定的距离，例如给定一个颜色，人很难判定其中的各个RGB分量，这时使用面向视觉感知的颜色空间比较方便。

在面向视觉感知的颜色空间中，人们通过将人类对颜色的观察情况组合而建立的孟塞尔颜色空间是一个基本的颜色空间，它主要基于主观观察，用于艺术领域。在此基础上，人们又提出了许多与孟塞尔系统近似的比较客观的颜色空间，如HIS、HSB等。这些颜色空间与人类关于颜色的视觉感知比较接近，同时独立于设备，是非线性的。

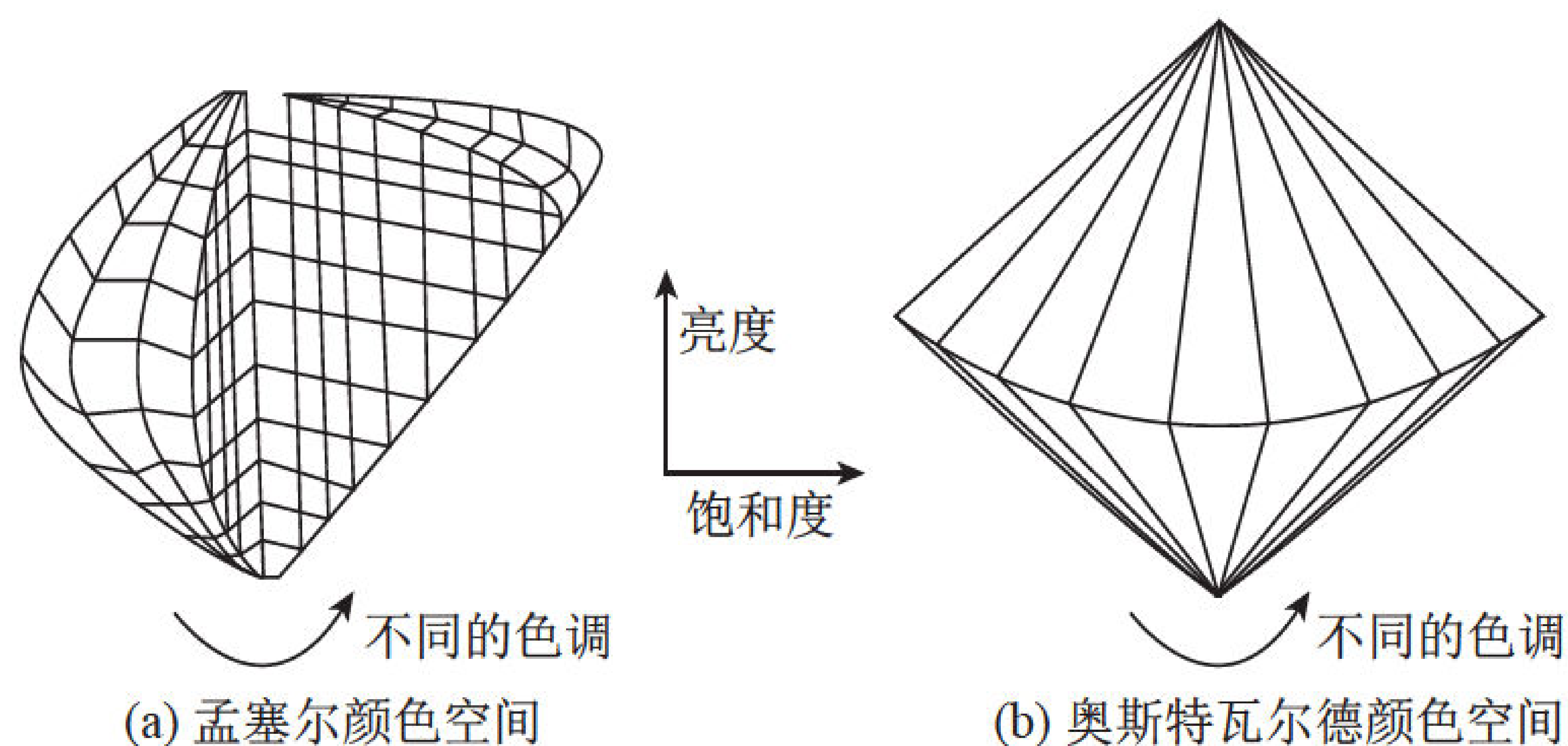
1. 由色卡系列构成的颜色空间

孟塞尔(Munsell)颜色空间以及奥斯特瓦尔德(Ostwalds)颜色空间是经过大量实验得到的一套实验数据或比色卡，不能直接应用于图像的表达和处理，但可以作为判断其他颜色空间独立性和均匀性的标准。孟塞尔和奥斯特瓦尔德颜色空间，都是色调、饱和度和亮度

的三维表达，它们都有沿着圆周的光谱色调。沿着垂直中线到颜色环平面亮度发生变化，主要的亮度轴被称为非彩色轴，沿着这个轴，颜色变化从白到灰，再到黑。饱和度表示为与非彩色轴之间的距离。孟塞尔颜色空间是目前使用的最重要的表色系统之一，美国和日本的颜色标准都以它为基础，并得到世界公认。

孟塞尔颜色空间于1905年被提出[149]，之后又经过几百万人次的实验和修正，其明度值以及色调、彩色的编排均接近视觉上等距的原则，因此常常作为判断其他颜色空间均匀性和独立性的标准。孟塞尔颜色空间包括1500种系统排列的颜色样本。这是一种均匀颜色空间的三维表示。一个色调环被分为5个主要色调，即红、黄、绿、蓝和紫。这些颜色又被进一步分为5个中间色调，即红黄、黄绿、绿蓝、蓝紫和红紫，而这些被分出来的颜色中的每一个又被划分为10种颜色，从而得到100种分离的色调。孟塞尔颜色空间的一个重要属性是饱和度和亮度的数量级别是按照相等的知觉间隔来确定的。[139]在三维空间中孟塞尔颜色空间为立体，如图4-2(a)所示。它的不规则外形是由于饱和度和亮度随着相等的知觉间隔而变化。

奥斯特瓦尔德颜色系统与孟塞尔颜色系统是相似的，但在奥斯特瓦尔德颜色系统中，定义每一种色调都是按固定比例数量混合得到的，8个相等空间分布的颜色从白到黑变化。如图4-2(b)所示，图中的这种规则外形是由于奥斯特瓦尔德颜色空间不依赖于相等的感知差别。



(a) 孟塞尔颜色空间　　(b) 奥斯特瓦尔德颜色空间

图4-2　孟塞尔和奥斯特瓦尔德颜色空间

2. XYZ颜色空间

由于用RGB颜色空间比配等能光谱时存在负比配，为了用三基色定义所有的颜色，国际照明委员会(CIE)定义了三种标准基色X、Y、Z，这三种基色是虚拟的，使颜色比配全

部为正值，称为XYZ颜色空间，是一种设备独立的颜色空间。在这个空间内，监视器的颜色范围成为一个变形的六面体。它与RGB颜色空间之间的关系是线性关系，其形状大约是一个顶点在坐标原点的圆锥体，锥体的外表面可以看成由一些从原点出发而终止于一条边的光线组成，这些光线可以看成某些色彩的各种颜色的集合，而Y值则表示某个颜色的亮度。

XYZ颜色空间包含所有人类能够感觉到的颜色，而且它是基于由实验测定的颜色匹配函数，因此它不同于RGB颜色空间只表示监视器所能显示的颜色范围，而是显示所有颜色。在XYZ颜色空间内，可以通过改变三个分量的数值来得到需要的颜色。

XYZ颜色空间的一个重要属性是它的设备非依赖性，每一个颜色空间都有一个来自CIE的XYZ颜色空间转变量。这个XYZ颜色空间通常作为一个参考颜色空间使用，它像一个中级非依赖装置颜色空间。

X、Y、Z和R、G、B系数之间的关系为

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 2.768\,9 & 1.751\,8 & 1.130\,2 \\ 1.000\,0 & 4.590\,7 & 0.060\,1 \\ 0.000\,0 & 0.056\,5 & 5.594\,3 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{4-10}$$

3. HSI颜色空间

HSI颜色空间是从人的心理感知角度建立的。H(色调)是指一种颜色在色谱中所对应的主波长，S(饱和度)相当于颜色的纯度，I(亮度)是感觉的均匀量。与HSI颜色空间相类似的还有HSV、HCV、HSB等颜色空间[16]，它们都属于极坐标空间结构，其中多数与RGB空间都是线性关系，因而是与设备相关的非线性颜色空间。这些颜色空间的优点是能直观地描述颜色。

HSI颜色空间有两个重要的特点。首先，亮度分量与色度分量是分开的，I分量与图像的彩色信息无关。其次，H及S分量与人感受彩色的方式紧密相连。这些特点使得HSI颜色空间非常适合基于人的颜色感知特性进行处理和分析的图像处理算法。其中，H分量对彩色描述的能力与人的视觉最接近，区分力比较强。[174]

HSI颜色空间的坐标系统接近圆柱坐标系统。对其中任意一个颜色点P，H值对应指向该点的向量与R轴的夹角。这个点的S与指向该点的向量长度成正比。I值与该点所在平面与最下方对应黑色点的距离成正比。如果颜色点在I轴上，则其S值为零而且H没有被定义，这些点被称为奇异点。奇异点的存在是HSI模型的一个缺陷，而且在奇异点附近，

RGB数值的微小变化会引起H、S、I数值的明显变化。

HSI颜色空间与RGB颜色空间可以通过一定的算法相互转换[112]，可使用下列方程

$$H=\begin{cases}\cos^{-1}\left(\dfrac{(R-G)+(R-B)}{2\sqrt{(R-G)^2+(R-B)(G-B)}}\right) & B\leqslant G\\ 2\pi-\cos^{-1}\left(\dfrac{(R-G)+(R-B)}{2\sqrt{(R-G)^2+(R-B)(G-B)}}\right) & B>G\end{cases} \tag{4-11}$$

$$S=1-\left(\frac{3}{R+G+B}\right)\min(R，G，B) \tag{4-12}$$

$$I=\frac{R+G+B}{3} \tag{4-13}$$

由式(4-11)直接计算出的H在[0，360]，为使H在[0，1]，可再令$H=H/360$进行转换。另外，当$S=0$时，无对应色，这时H没有意义，此时定义H为0；当$I=0$或$I=1$时，S值没有意义。

处理图像时常需要将HSI转换为RGB。在[0，1]给出HSI值，并在相同的值域找到RGB值，可利用H值公式。在原始分割中有3个相隔120°的扇形，如图4-3所示。从H乘以360°开始，这时色调值返回到原来的范围[0°，360°]。

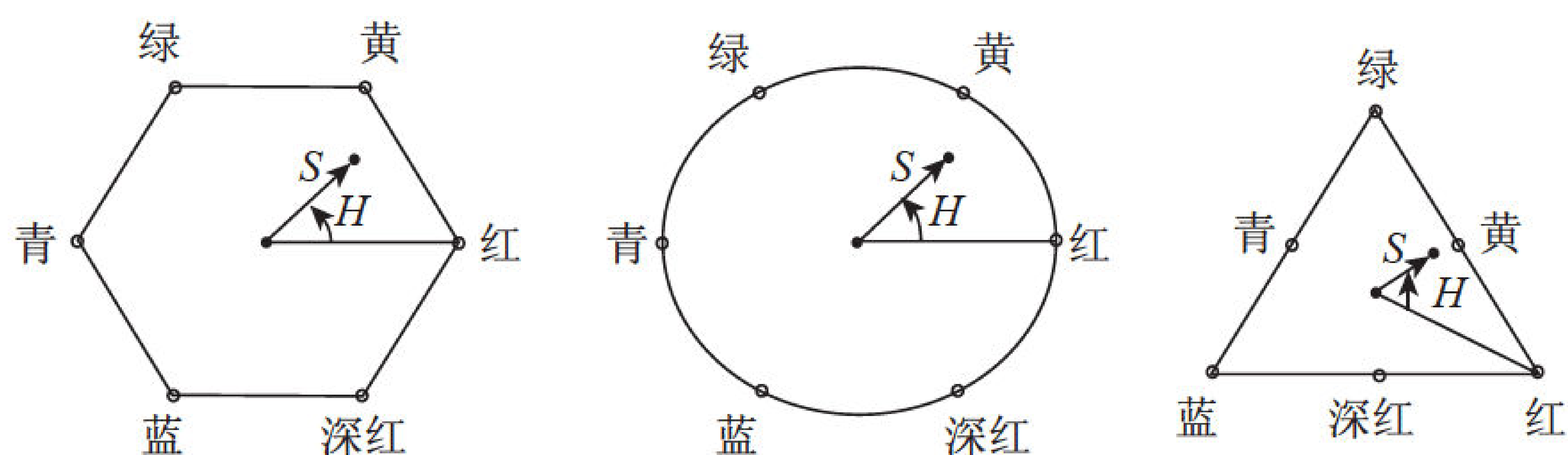


图4-3　HSI彩色空间中的色调和色饱和度

RG扇区(0°≤H<120°)：当H位于这一扇区时，RGB分量为

$$B=I(1\square S) \tag{4-14}$$

$$R=I\left[1+\frac{S\cos H}{\cos(60^\circ-H)}\right] \tag{4-15}$$

$$G=1\square(R+B) \tag{4-16}$$

GB扇区($120^\circ \leqslant H < 240^\circ$)：如果给定的$H$值在这一扇区，首先从$H$中减去120°，即

$$H^*=H-120^\circ \tag{4-17}$$

然后RGB分量为

$$R=I(1-S) \tag{4-18}$$

$$G=I\left[1+\frac{S\cos H^*}{\cos\left(60^\circ-H^*\right)}\right] \tag{4-19}$$

$$B=1-(R+G) \tag{4-20}$$

BR扇区($240^\circ \leqslant H < 360^\circ$)：如果$H$在这一扇区，从$H$中减去240°，即

$$H^*=H-240^\circ \tag{4-21}$$

然后RGB分量为

$$G=I(1-S) \tag{4-22}$$

$$B=I\left[1+\frac{S\cos H^*}{\cos\left(60^\circ-H^*\right)}\right] \tag{4-23}$$

$$R=1-(G+B) \tag{4-24}$$

HSI颜色空间提供了能更好地符合人眼颜色感知习惯的描述颜色的三个分量，但RGB仍然存在色觉差异非线性的缺点，尤其在H方向上的颜色与角度是非线性的。[135]

4.1.3　均匀颜色空间

均匀颜色空间本质上仍是面向视觉感知的颜色空间，只是在视觉感知方面更为均匀。在对颜色的感知、分类和鉴别中，对颜色的描述应该是越准确越好。从图像处理的角度来看，对颜色的描述应该与人对颜色的感知越接近越好。从视觉均匀的角度来看，人所感知到的两种颜色的距离最好与这两个颜色在表达它们的颜色空间中的距离成比例。[175]

如果在颜色空间中任选一点，通过该点的任意方向上相同的距离能表示相同的颜色感觉变化，则称这样的颜色空间为均匀颜色空间。均匀颜色空间的确定试图完全按照人类对颜色的感知来划分颜色。[23]

1. CIE$L^*a^*b^*$颜色空间

CIE$L^*a^*b^*$颜色空间是为感觉均匀而设计的一个国际标准，其目的是使人感受的颜色差

别等同对应CIE$L^*a^*b^*$颜色空间中相等的欧几里得距离，以及所有颜色都按照其实验测得的相互之间知觉色差的多少，尽可能均匀地分布于颜色空间，这是目前最均匀的颜色空间。

CIE早期推荐的颜色空间主要有麦克斯韦三角形、RGB坐标、XYZ坐标等。因为提出它们的最初目的只是充分而方便地表达自然界的色彩，没有考虑人眼的颜色分辨特性，所以独立性和均匀性都较差。如果颜色模型能提供空间位置与视觉系统之间良好的对应关系，就会给颜色研究工作带来很多便利。

CIE于1976年提出的CIE$L^*a^*b^*$色度系统就是这样一个颜色空间，其中L^*、a^*、b^*坐标分别表示明度、红和黄，空间中的位置指明其对应的颜色。在该系统中，明度L表示颜色明亮的程度；a^*表示红色在颜色中占有的成分，□a^*表示红色的补色在颜色中占有的成分；b^*代表颜色中黄色的成分，□b^*表示黄色的补色在颜色中占有的成分。CIE$L^*a^*b^*$颜色空间是基于对立色理论和参考白点而建立的[120]，与设备无关，适用于接近自然光照的应用场合。

三个刺激量X、Y、Z到CIE$L^*a^*b^*$颜色空间的转换公式为

$$L^* = \begin{cases} 116f(Y/Y_0)^{\frac{1}{3}} - 16 & \text{当} Y/Y_0 > 0.008\,856 \\ 903.3f(Y/Y_0)^{\frac{1}{3}} & \text{当} Y/Y_0 \leqslant 0.008\,856 \end{cases} \tag{4-25}$$

$$a^* = 500\left[f(X/X_0) - f(Y/Y_0)\right] \tag{4-26}$$

$$b^* = 200\left[f(Y/Y_0) - f(Z/Z_0)\right] \tag{4-27}$$

其中

$$f(t) = \begin{cases} t^{1/3} & \text{当} t > 0.008\,856 \\ 7.787t + 116 & \text{当} t \leqslant 0.008\,856 \end{cases} \tag{4-28}$$

各式中的下标0表示对应的参考白点。

CIE$L^*a^*b^*$强调对绿色的表示(对绿色比较敏感)，其次是红色和蓝色。

2. CIE$L^*u^*v^*$颜色空间

CIE同时定义了CIE$L^*u^*v^*$颜色空间，与CIE$L^*a^*b^*$类似，u^*、v^*表示由色调和彩度形成的色感知属性。CIE$L^*u^*v^*$和CIE$L^*a^*b^*$有一些共同的特点。首先，它们的颜色空间虽然不能达到完全均匀，但基本上是均匀的。其次，因为基于颜色测量的CIE系统，CIE$L^*u^*v^*$和CIE$L^*a^*b^*$都是与设备无关的颜色空间。

CIE$L^*u^*v^*$颜色空间同样是基于对立色理论和参考白点而建立的[120]，与设备无关，适用于显示器显示和根据加色原理进行组合的应用场合。[176]

三个刺激量X、Y、Z到CIE$L^*u^*v^*$颜色空间的转换公式为

$$L^* = \begin{cases} 116(Y/Y_0)^{\frac{1}{3}} - 16 & \text{当} Y/Y_0 > 0.008\,856 \\ 903.3(Y/Y_0)^{\frac{1}{3}} & \text{当} Y/Y_0 \leqslant 0.008\,856 \end{cases} \tag{4-29}$$

$$u^* = 13L(u' - u_0') \tag{4-30}$$

$$v^* = 13L(v' - v_0') \tag{4-31}$$

变量u'、v'和u'_0、v'_0可以从下式得出

$$u' = \frac{4X}{X + 15Y + 3Z} \tag{4-32}$$

$$u'_0 = \frac{4X_0}{X_0 + 15Y_0 + 3Z_0} \tag{4-33}$$

$$v' = \frac{9Y}{X + 15Y + 3Z} \tag{4-34}$$

$$v'_0 = \frac{9Y_0}{X_0 + 15Y_0 + 3Z_0} \tag{4-35}$$

X_0、Y_0、Z_0为标准白光所对应的X、Y、Z值。CIE$L^*u^*v^*$颜色空间强调对红色的表示，即对红色变化比较敏感，但对蓝色表示得比较粗糙。

CIE$L^*u^*v^*$和CIE$L^*a^*b^*$颜色空间是为了解决颜色的等距度量而建立的，具有均匀明度标尺和均匀色品标尺，其独立性和均匀性显然优于XYZ空间，但是，由于均匀色品标尺没有把色调和饱和度分离开，所以只有一个亮度分量能用于处理，而不能利用彩色图像的色彩信息，这会把彩色图像处理限制在灰度图像处理的范围，极大地束缚了彩色图像处理的手脚，不是应选的最佳颜色空间。相对来讲，它是目前推荐的最好用的量化分析颜色的均匀空间。除了继承XYZ颜色空间的优点外，它还有两个新的优点：①颜色差异的规范化空间；②CIE$L^*u^*v^*$和CIE$L^*a^*b^*$颜色空间可以由圆柱坐标表示，因此可以满足不同的需要。由于具有这些优点使这两种颜色空间在对颜色有严格要求的应用场合中更加实用。

3. LHS颜色空间

在CIE$L^*u^*v^*$空间中，以等能白光的色品坐标(u_0，v_0)为中心，如果其他色品位置(u，v)与之连线，那么连线的长度可以代表饱和度，连线相对于某一基准线的角度可以

代表色调。因此，为了克服CIE颜色空间的固有缺陷，有人从Luv空间中推导出亮度－色度－饱和度(即LHS)颜色空间。其表达式为

$$
\begin{aligned}
L &= L \\
H &= \arctan\left(\frac{v-v_0}{u-u_0}\right) \\
S &= [(u-u_0)^2+(v-v_0)^2]^{1/2}
\end{aligned}
\tag{4-36}
$$

因为LHS空间的三个分量能在较大的范围内满足彩色图像处理的两个要求，而且它们与人眼的颜色感觉相对应，可以有目的地进行处理，因而在彩色图像处理中具有比其他颜色空间更好的优越性。

该颜色空间具有两个缺陷：一是在处理过程中如果某一分量的变化幅度超过一定的范围，会引起其他分量的变化，即独立性不是很好。例如，亮度分量改变较大时，可能会引起色调的变化。二是亮度分量和饱和度分量存在“越界”问题(即超过允许的最大值)。

4.1.4 色差度量

欧氏距离内颜色距离度量是最简单的度量方法之一。在RGB图像中，给定一个有代表性的彩色点样品集，可得到一个彩色“平均”估计，用RGB向量a来表示这个平均彩色。令z代表RGB空间中的任意一点，如果它们之间的距离小于特定的阀值D_0，我们说z和a是相似的，z和a之间的欧氏距离为

$$
\begin{aligned}
D(z，a) &= \|z-a\| = [(z-a)^T(z-a)]^{\frac{1}{2}} \\
&= (z_R-a_R)^2+(z_G-a_G)^2+(z_B-a_B)^2]^{\frac{1}{2}}
\end{aligned}
\tag{4-37}
$$

其中，注脚R、G、B表示向量a和z的RGB分量。$D(z，a)\leqslant D_0$的点的轨道是半径为D_0的实心球。包含在球内部和表面上的点符合特定的彩色准则；球外面上的点则不符合准则。

公式(4-37)一个有用的推广是如下形式的距离测度

$$
D(z，a) = [(z-a)^T C^{-1}(z-a)]^{\frac{1}{2}} \tag{4-38}
$$

这里C是希望分割的彩色典型样本的斜方差矩阵。因为距离是正的和单调的，可以用距离的平方运算来代替，这样就避免了开方运算。然而即使不计算平方根，执行式(4-38)

的计算代价也很高。折中方案是使用边界盒。在该方法中，盒的中心在a上，沿每一彩色轴的尺度选择与沿每一个轴取样的标准差成比例，标准差的计算只使用一次样本数据。

RGB颜色空间内的欧式距离计算中没有充分考虑颜色的视觉感觉特性，实际中有时会产生错误的色差度量结果，因此可以考虑在均匀颜色空间内的色差度量。

在均匀颜色空间内进行色差的度量是相对理想的方式，例如在CIE$L^*u^*v^*$颜色空间内色差公式为

$$\Delta E_{uv}=\sqrt{\left(\Delta L^*\right)^2+\left(\Delta u^*\right)^2+\left(\Delta v^*\right)^2} \tag{4-39}$$

其中，$\Delta L^*=L^*{}_2\square L^*{}_1$，$\Delta u^*=u^*{}_2\square u^*{}_1$，$\Delta v^*=v^*{}_2\square v^*{}_1$。

CIE$L^*a^*b^*$颜色空间内色差公式为

$$\Delta E_{ab}=\sqrt{\left(\Delta L^*\right)^2+\left(\Delta a^*\right)^2+\left(\Delta b^*\right)^2} \tag{4-40}$$

其中，$\Delta L^*=L^*{}_2\square L^*{}_1$，$\Delta a^*=a^*{}_2\square a^*{}_1$，$\Delta b^*=b^*{}_2\square b^*{}_1$。

上文给出了用于颜色色差度量的一些基本方法。由于RGB颜色空间本身的不均匀性，它并不能很好地满足颜色视觉特性，而采用其他近似均匀的颜色空间，存在空间变换复杂、计算烦琐等缺陷。如何得到既可以适度地满足颜色视觉特性，又计算简单，是彩色图像处理中色差计算所追求的目标。

4.2　颜色相似性度量

色差度量本质上是颜色之间相似性的度量，色差大代表颜色相似性程度小，反之，色差小则代表颜色相似性程度大。因此，我们可以从另外一个角度，即利用颜色相似性来分析两种颜色之间的色差。颜色相似性度量在彩色图像边缘检测、区域分割以及图像检索、图像压缩等领域中有广泛的应用，是其他彩色图像研究领域的基础，起到重要的作用。

颜色相似性度量与颜色的表示方式——颜色空间有密切的关系。在均匀颜色空间内，实现颜色相似性度量，将更加符合人眼颜色视觉特性，得到效果更佳的彩色处理结果。对

各类颜色空间的定义、转换以及色差度量的分析结果表明，目前还没有设计出完全满足人眼视觉特性的颜色空间。相对来讲，CIE$L^*u^*v^*$和CIE$L^*a^*b^*$以及HSI等颜色空间更加符合颜色视觉特性的要求。彩色图像处理中经常采用HSI颜色空间表示颜色，实现色差度量，可以在一定程度上满足颜色视觉特性的要求。

彩色图像最常用的颜色表示是RGB颜色空间，是一种基本的颜色描述方法，但由于RGB颜色空间的非均匀性，直接利用这些分量进行彩色图像处理常常不能得到所需的效果，需要选择能够较好地表示彩色图像的颜色空间。实际中，常将RGB颜色空间转换到其他颜色空间中去。但由于在RGB颜色空间变换过程中计算复杂，这影响了大数据量彩色图像处理的实时性要求。

颜色相似度量是彩色图像处理中的基本运算，要求在一定程度上满足颜色视觉特性，同时要保证一定的实时性。一种理想的颜色相似度量方法应该采用基本的RGB颜色空间表示颜色，不必进行复杂的颜色转换过程，同时满足颜色视觉特性。据此，本节在前述研究的基础上提出了一种新的颜色相似性度量方法，并给予具体介绍。

4.2.1 RGB与HSI颜色空间变换的特征

设彩色图像中的一个像素为P，在RGB颜色空间内表示为$P=[R，G，B]'$，从RGB到HSI的颜色空间转换可参照公式(4-11)、公式(4-12)、公式(4-13)进行计算。

设彩色图像中的另一个像素为P'，在RGB颜色空间内表示为$P'=[KR，KG，KB]'$，这里K为常值比例系数。那么关于P'的色度H'、饱和度S'、亮度I'的值分别为(以$B \leqslant G$为例，$B>G$类同)

$$\begin{aligned} S' &= 1-\frac{3}{KR+KG+KB}\left[\min(KR，KG，KB)\right] \\ &= 1-\frac{3}{K(R+G+B)}\left[K\min(R，G，B)\right] \\ &= 1-\frac{3}{R+G+B}\left[\min(R，G，B)\right] \\ &= S \end{aligned} \tag{4-41}$$

$$
\begin{aligned}
H' &= \cos^{-1}\left\{\frac{(KR-KG)+(KR-KB)}{2\sqrt{(KR-KG)^2+(KR-KB)(KG-KB)}}\right\} \\
&= \cos^{-1}\left\{\frac{K(R-G)+K(R-B)}{2K\sqrt{(R-G)^2+(R-B)(G-B)}}\right\} \\
&= \cos^{-1}\left\{\frac{(R-G)+(R-B)}{2\sqrt{(R-G)^2+(R-B)(G-B)}}\right\} \\
&= H
\end{aligned} \tag{4-42}
$$

$$
I' = \frac{KR+KG+KB}{3} = K\frac{R+G+B}{3} = KI \tag{4-43}
$$

综上所述，如果两种颜色RGB值的大小成比例，则它们的色度和饱和度值是相等的，亮度值成K倍关系。这是在进行RGB与HSI颜色空间变换过程中表现出的一个特征，利用这个特征我们可以进一步提出进行颜色相似性度量的新方法。

4.2.2　颜色相似系数

1. 色度饱和度相似系数

在RGB颜色空间中，颜色向量是用三维向量来表示的。对于三维空间内的一个向量来说，可由两个参量决定，即向量的方向和幅值。判断两个向量之间的相似性，也由这两个参量来决定。两个向量之间的方向相似性度量可由相关系数的计算得到。

设RGB颜色空间中的两种颜色用颜色向量分别表示为f和g，这里$f=(f_1, f_2, f_3)'$，$g=(g_1, g_2, g_3)'$。这两个颜色向量之间的相关系数用$r(f, g)$表示，则$r(f, g)$定义为

$$
r_{(f,\ g)} = \frac{\langle f,\ g\rangle}{\|f\|\|g\|} \tag{4-44}
$$

其中，$\langle f,\ g\rangle = f'g$，$\|f\| = \sqrt{(f,\ f)}$，$\|g\| = \sqrt{(g,\ g)}$，$r(f,\ g) \in [0,\ 1]$。

当两个颜色向量方向相同时，$f=Kg$(K为常数)，即$f_1=Kg_1$，$f_2=Kg_2$，$f_3=Kg_3$，则$r(f, g)=1$。随着两个颜色向量之间的夹角增大，$r(f, g)$值减小。这样可以通过两种颜色在RGB颜色空间内计算得到相关系数$r(f, g)$来表示在HSI颜色空间中色度H和饱和度S的相似程度。$r(f, g)$值越大，相似程度越大，反之亦然。这个相关系数被称为色度饱和度相似系数，可直接用于颜色相似性的分析计算中。

2. 亮度相似系数

两种颜色的相似系数为1，只是代表颜色各个分量成一定比例，即HSI空间中对应的色度和饱和度是相等的，但并不代表对应的亮度值相等，它们的亮度有可能相差很大，如果这样就不能把两种颜色划入一个颜色一致性区域中。因此只有一个色度饱和度相似系数还无法完全表示出两个像素的颜色相似性程度，需要进一步比较颜色的亮度。因此，根据RGB与HSI颜色空间变换公式，可定义一个比较两种颜色f和g之间的亮度相似系数$k(f, g)$，即

$$k_{(f,g)}=1-\frac{\left|f_1+f_2+f_3-g_1-g_2-g_3\right|}{C} \tag{4-45}$$

这里$C=3\times 255=765$。$k(f, g)\in[0, 1]$，当两个亮度相同时，系数$k(f, g)=1$，当两个亮度差加大时，$k(f, g)$值减小。

3. 颜色相似系数

从人眼颜色视觉特性来分析，色度和饱和度的区别是主要的方面，决定了颜色的主要特性。所以在上述的两个系数中，色度饱和度相似系数代表了颜色相似性的主要方面，而亮度相似系数代表颜色相似性的次要方面。综合考虑这两个系数，可以分别对两个系数加权，并设定权值来调整两个系数在颜色相似性度量中的重要程度，这样可以得到一个综合的颜色相似性度量系数$s(f, g)$，简称颜色相似系数，即

$$s(f, g)=\alpha_1 r(f, g)+\alpha_2 k(f, g) \tag{4-46}$$

其中$\alpha_1+\alpha_2=1$，$\alpha_1>0$，$\alpha_2>0$。

颜色相似性本身具有一定的模糊性，颜色相似与否不能简单以{0，1}来表示，而应以连续变化的数值范围表示两种颜色之间的相似程度。由于颜色相似系数的数值范围在[0，1]，其间不同数值的大小表示两种颜色之间的相似性大小，很好地满足了人眼对于颜色之间相似性的感知属性。以颜色相似系数为基础的颜色相似度量方法，可广泛用于彩色图像处理。

4. 颜色相似性判决

定义了颜色相似系数，就可以判断两种颜色是否相似。根据颜色相似系数的大小，针对不同的应用目的，确定相应的颜色相似性判定阈值大小，判别两种颜色是否相似。判定两种颜色是否相似的规则为

$$\begin{cases} s(f, g) \geq Ts，两个颜色是相似的 \\ s(f, g) < Ts，两个颜色是不相似的 \end{cases} \tag{4-47}$$

这里的Ts是颜色相似性判定阈值。

4.3　实验结果与分析

选取12种颜色作为样本颜色，如图4-4所示。按从上到下、从左到右的顺序，依次为各种颜色编排序号为1－12。12种颜色的RGB数值及颜色感觉，如表4-1所示。选取第12号颜色样本即深红色作为基准样本。按照计算它与其他颜色样本之间的色度饱和度相似系数、亮度相似系数以及颜色相似系数，最后根据颜色相似系数的大小判断12号颜色样本与其他颜色样本之间的颜色是否满足相似关系，数据结果如表4-2所示。

实验中，取α_1=0.85，α_2=0.15，Ts=0.98。表4-2中的数据显示，颜色相似系数的定义和计算结果在一定程度上满足人眼颜色视觉的相似性判定规律。



图4-4　12种颜色样本

表4-1　12种颜色样本的RGB值

颜色序号	颜色感觉	RGB值		
		R	G	B
1	红	240	16	16
2	绿	16	240	16
3	蓝	16	16	240
4	黄	240	240	16
5	紫	144	16	240
6	橙	248	48	36

(续表)

颜色序号	颜色感觉	RGB值		
		R	G	B
7	粉	208	16	240
8	棕	134	16	16
9	灰	128	128	128
10	白	240	240	240
11	黑	16	16	16
12	深红	216	8	24

表4-2　颜色相似计算结果

颜色序号	颜色感觉	色度饱和度相似系数	亮度相似系数	颜色相似系数	是否相似
1	红	0.998 3	0.947 7	0.985 6	是
2	绿	0.113 3	0.947 7	0.405 3	否
3	蓝	0.187 0	0.947 7	0.453 3	否
4	黄	0.734 4	0.654 9	0.706 6	否
5	紫	0.613 8	0.780 4	0.672 1	否
6	橙	0.984 2	0.884 1	0.966 1	否
7	粉	0.740 8	0.696 7	0.725 4	否
8	棕	0.968 9	0.885 0	0.946 3	否
9	灰	0.664 4	0.801 3	0.712 3	否
10	白	0.664 4	0.362 1	0.558 6	否
11	黑	0.664 4	0.759 5	0.697 7	否

4.4　结论

通过上述分析得出以下结论：

(1) 本章对常用且具有代表性的几种颜色空间进行了分类，并分别给予相对详细的介绍；对它们的设备相关性、感知均匀程度、视觉直观性以及色差度量等方面进行了分析和讨论。

(2) RGB颜色空间是最常用的一种颜色空间，大多用于彩色图像的输入和输出。但由于RGB颜色空间本身的缺陷，在彩色图像处理中经常需要把RGB颜色空间转换到其他颜色空间进行处理。颜色空间转换是研究设备彩色特性和进行色彩管理工作的基础。本章对RGB颜色空间与其他各类颜色空间的转换方法进行了详细的介绍。

(3) 一个独立性和均匀性较好的颜色空间将会使彩色图像处理发挥更大的潜力，是彩色图像处理中必须首先考虑的问题。但目前还没有完全满足人眼颜色视觉特性的颜色空间存在，如何得到完全均匀的颜色空间是需要进一步解决的问题。

(4) HSI颜色空间是相对符合人眼颜色视觉特性的颜色空间，在彩色图像处理中得到广泛应用。但在颜色空间变换过程中存在复杂的计算，影响了彩色图像处理的实时性。

(5) 根据HSI与RGB颜色空间的转换公式，推导出两种空间各颜色分量之间的数值对应关系。据此，定义了色度饱和度相似系数、亮度相似系数。根据人眼在分辨颜色时的特性，最终定义颜色相似系数，确定了两种颜色相似的判别规则。通过实验证明了该方法符合人眼颜色相似判别规律。该方法计算简单，易于实现，可用于彩色图像匹配中的像素颜色相似性计算。

第5章 彩色遥感图像拼接



彩色图像由于具有更为丰富的可视化信息，因此得到了广泛的应用。彩色图像中像素颜色通常需要由三维颜色空间向量来表示，因此与灰度图像相比，在图像的表达方式和处理方法上存在较大差异，使得彩色图像拼接具有更大的难度。以往的彩色图像拼接方法大多首先将彩色图像变换为灰度图像，然后采用灰度图像拼接的方法进行匹配计算。由于在彩色与灰度图像变换过程中，必然会丢失大量有用的颜色信息，因此难以得到理想的拼接处理结果。彩色图像平滑过程中可以通过对图像的拼接缝做各种颜色的过渡处理，但由于图像整体上存在颜色差异，使平滑后的图像很难保证整体上的颜色一致性。[7][8][9]人眼对于彩色具有较高的敏感程度，彩色图像平滑同样成为图像拼接中的另一个关键技术环节。由于彩色图像的数据量更大，如何提高拼接算法的效率是需要特别考虑的问题。

基于第4章对RGB与HSI颜色空间特性及其变换特征的分析和颜色相似性判断，本章通过引入协方差矩阵运算，实现彩色图像特征模板的提取。通过最小二乘法和伽马校正方法，建立两幅图像之间的亮度变换函数，实现彩色图像平滑计算。

5.1 颜色空间变换

颜色空间是颜色的空间表达形式，包括多种类型，如RGB、HSI和YUV等。[64]颜色空间把抽象的、主观的视觉感受具体化为三维空间中的位置或向量，使各种彩色图像、彩色设备的颜色特性都能实现可视化，成为颜色研究的重要工具。

由于图像采集和显示设备较多使用的是RGB颜色空间，所以RGB颜色空间成为彩色图像处理中最基础、最常用的颜色空间。然而，RGB颜色空间存在固有的缺点：①不直观，即从RGB值中很难看出其所表示的颜色的认知属性；②不均匀，即两个色点之间的距离不等于两个颜色之间的知觉差异；③对硬件设备具有依赖性。因此，RGB颜色空间是一个与设备相关的、颜色描述不完全直观的空间。为了克服RGB颜色空间的不均匀和不直观的缺点，在彩色图像处理中大多采用颜色空间变换方法，将RGB颜色空间转换为其他更符合颜色视觉特性的颜色空间，再做进一步的处理，并将处理后的结果重新转换为RGB颜色空间。颜色空间的转换过程需要进行大量的计算，对彩色图像处理的实时性

有较大的影响。

另一种常用的颜色空间是HSI颜色空间，其中，H(色调)表示一种颜色在色谱中对应的主波长，S(饱和度)代表颜色的纯度，I(亮度)代表感觉的均匀量。[16]HSI颜色空间有两个重要的特点：①亮度分量与色彩分量是分开的，I分量与图像的彩色信息无关；②H及S分量与人感受彩色的方式紧密相连。从人的心理视觉和生理视觉角度来看，HSI颜色空间对于颜色的描述显然更符合人的视觉特性，因此在较多的图像处理应用中，须将RGB颜色空间转换为HIS颜色空间表达后再进行处理，以得到更佳的彩色图像处理效果。

HSI颜色空间与RGB颜色空间可以通过公式(4-11)、公式(4-12)、公式(4-13)实现转换[112]，如果两种颜色的RGB值(r，g，b)(r'，g'，b')大小均以K倍比例发生变化，即

$$\frac{r}{r'}=\frac{g}{g'}=\frac{b}{b'}=K \tag{5-1}$$

则它们的色度和饱和度值是相等的，但亮度值将为K倍关系。这是在进行RGB与HSI颜色空间变换过程中表现出的一个重要特征，根据这个特征可以进一步提出用于彩色图像拼接的匹配计算的新方法。

5.2　色差计算

在满足等式(5-1)的条件下，可以得到以下三个等式

$$r*g'=g*r' \tag{5-2}$$

$$r*b'=b*r' \tag{5-3}$$

$$b*g'=g*b' \tag{5-4}$$

上述三个等式中，只要有两个等式成立，则另外一个等式一定成立。考虑到计算量，可以给出以下颜色色差度量公式

$$E_{rg}=|r*g'-g*r'| \tag{5-5}$$

$$E_{bg}=|b*g'-g*b'| \tag{5-6}$$

在E_{rg}和E_{bg}均较小的条件下，可判定两种颜色的色度和饱和度是近似相等的。由于色度、饱和度和亮度三个颜色分量中，色度和饱和度决定了颜色的主要类别，可取以上两项

的和值作为度量两种颜色色差大小的变量。

$$E=E_{rg}+E_{bg} \tag{5-7}$$

5.3 彩色图像匹配

5.3.1 特征模板提取

图像拼接要求待拼接的相邻两幅图像边界必须存在重叠区域，重叠区域具有相同的景物内容。特征模板是指在一幅图像的重叠区域中搜索到的、可进行有效图像匹配计算的图像区域。一般来说，图像模板的选择要满足两个原则：选择特征模板的尺寸不宜过大，尽管大尺寸的特征模板可提高匹配计算的精度，但会使匹配过程所需的计算量急剧增加；另外，要求特征模板中的图像灰度或颜色信息较为丰富，以保证匹配计算的准确性。因此，特征模板的选择在图像匹配计算中很重要。对于彩色图像来说，合理的特征模板所在的图像区域应包含更多类别的颜色，具有更多的图像纹理或边缘。具体来说，要求特征模板中的像素颜色在RGB颜色空间内较离散地分布，颜色与颜色之间具有较小的相关性。为了在图像重叠区域中，搜索到颜色分布合理的特征模板，这里引入多维数据的数字特征分析及相关矩阵计算，以实现有效的后续匹配运算。

令$(X_1, X_2, X_3)^T$为从彩色图像模板区域中所获取的RGB颜色空间表示的三维样本数据集合：$\{(X_{11}, X_{12}, X_{13})^T, (X_{21}, X_{22}, X_{23})^T, \cdots, (X_{n1}, X_{n2}, X_{n3})^T\}$这里的$n$表示颜色样本的个数，即特征模板图像区域中像素的个数。记各组颜色样本数据为

$$x_i=(x_{i1}, x_{i2}, x_{i3})^T, \quad i=1, 2, \cdots n \tag{5-8}$$

其中，X_1代表样本数据的R颜色分量矢量，X_2代表样本数据的G颜色分量矢量，X_3代表样本数据的B颜色分量矢量。引进$n\times3$颜色样本数据观测矩阵为

$$X=\begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & x_{n3} \end{bmatrix}=\begin{bmatrix} x_1^T \\ x_2^T \\ \vdots \\ x_n^T \end{bmatrix} \tag{5-9}$$

X的3个列分别是3个矢量X_1、X_2、X_3的n个观测数据。其中，第j列数据的均值为

$$\overline{x}_j=\frac{1}{n}\sum_{i=1}^{n}x_{ij}，\ j=1，2，3 \tag{5-10}$$

第j列数据的方差为

$$s_j^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_{ij}-\overline{x}_j)^2，\ j=1，2，3 \tag{5-11}$$

第j、k列数据的协方差为

$$s_{jk}=\frac{1}{n-1}\sum_{i=1}^{n}(x_{ij}-\overline{x}_j)(x_{ik}-\overline{x}_k)，\ j、k=1，2，3 \tag{5-12}$$

并且有

$$S_j^2=S_{jj}，\ j=1，2，3 \tag{5-13}$$

称$\overline{x}=(\overline{x}_1，\overline{x}_2，\overline{x}_3)^{\mathrm{T}}$为样本观测数据的均值向量。令

$$S=\begin{bmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{bmatrix} \tag{5-14}$$

则S为颜色样本观测数据的协方差矩阵。X的第$(j，k)$列数据的相关系数为

$$\rho_{jk}=\frac{s_{jk}}{\sqrt{s_{jj}}\sqrt{s_{kk}}}=\frac{s_{jk}}{s_j s_k}，\ j、k=1，2，3 \tag{5-15}$$

其中，$\rho_{jj}=1$，$|\rho_{jk}|\leqslant 1$。令

$$P=\begin{bmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{21} & 1 & \rho_{23} \\ \rho_{31} & \rho_{32} & 1 \end{bmatrix} \tag{5-16}$$

表示观测数据的Pearson相关矩阵。相关矩阵P表示变量观测值之间线性相关的密切程度。由于矩阵元素的对称性，这里只需考虑相关矩阵P的下三角矩阵元素。为了确定采样得到的彩色图像特征模板是否具有丰富的像素颜色数据分布状态，可采用

$$\lambda=\rho_{21}\times\rho_{31}\times\rho_{32} \tag{5-17}$$

来确定，λ值越小，表示颜色样本空间具有各个像素的颜色向量的相关性越小，相应地，颜色的色度和饱和度差异越大，因此更适合选作特征模板。一般情况下，图像中灰度或颜色变化具有相似性，为提高特征模板的搜索速度，减少计算量，在特征模板搜索过程中，可采用跳跃式搜索方式来得到最佳的图像特征模板。

5.3.2 匹配搜索策略

以特征模板为基准，在待拼接的图像重叠区域中，采用本书5.2节中提出的颜色色差度量方法进行匹配计算，记录所有匹配位置处的模板总体颜色相似度量数值，取其中最小的匹配位置为最佳匹配位置，具体计算过程如下所述。

设待拼接的两幅图像为I_1和I_2，两幅图像的重叠区域分别设为O_1和O_2。在图像区域O_1中，根据上文提出的特征模板提取方法，搜索得到特征模板T，其尺寸大小为$M\times N$。在图像区域O_2中，按从上到下、从左到右的顺序，移动特征模板T，与具有相同尺寸大小的图像区域Y进行匹配计算，得到

$$\mathrm{sum}(m,\ n)=\sum_{i=1}^{M}\sum_{j=1}^{N}E_{(i,\ j)} \tag{5-18}$$

在完成所有的匹配值计算后，取其中的最小值所对应的重叠区域内的相对位置为最佳匹配位置

$$(m_0,\ n_0)=\min[\mathrm{sum}(a,\ b)],\ (a,\ b)\in O_2 \tag{5-19}$$

5.4 彩色图像平滑

由于人眼对于彩色信息更为敏感，彩色图像拼接中拼接缝或图像整体颜色的不协调将使图像的视觉效果受到影响，因此彩色图像拼接中图像颜色平滑的问题更为突出。常用的图像颜色平滑方法包括基于图像直方图、基于图像信息熵、基于相邻图像方差、均值的色调调整方法等。这些方法都是以相邻图像之间的重叠区域或相邻图像之间拼接缝为基础，通过颜色的平滑过渡来消除人为的假边缘和拼接缝。在相邻图像重叠区域较大的情况下，对于两幅相邻图像拼接缝的消除都有一定的效果，但无法真正消除大幅面图像的颜色差异。从根本上解决这类问题的方法应是以一幅图像为基准，通过建立序列图像之间的颜色变换关系，从全局调整其他整幅图像的颜色分布，使合成的大幅面图像的颜色协调一致。前述的匹配运算是以色度和饱和度为标准来实现的，可保证匹配后的两幅图像所对应的像素色度和饱和度的一致性，但无法保证其亮度的一致性，因此需要对在亮度上可能存在的

差异进行调整，从而实现两幅图像颜色的一致性。本书将以最小二乘法为数学工具，通过引入伽马校正，建立图像之间的亮度变换函数，实现图像的亮度平滑。

5.4.1　伽马校正方法

图像增强处理中，幂次变换是根据幂次规律对图像亮度进行调整，从而得到新的响应输出。幂次变换在图像获取、打印和显示的各种装置中得到了较好的应用。幂次变换的基本形式为

$$y=\beta x^{\gamma} \tag{5-20}$$

其中β和γ为正的常数。

作为x的函数，y对于γ的各种值绘制的曲线，如图5-1所示($\beta=1$，γ值从上到下依次为：0.04，0.10，0.20，0.40，0.67，1，1.5，2.5，5.0，10.0，25.0)。从图5-1可以看出，随着γ值的变化将得到一组变换曲线。通过γ值把输入窄带暗值映射到宽待输出值。相反，输入高值时也成立，$\gamma>1$的值和$\gamma<1$的值产生的曲线有相反的数值效果。习惯上，将幂次变换中的指数称为伽马值，而把修正幂次函数响应现象的过程称为伽马校正。

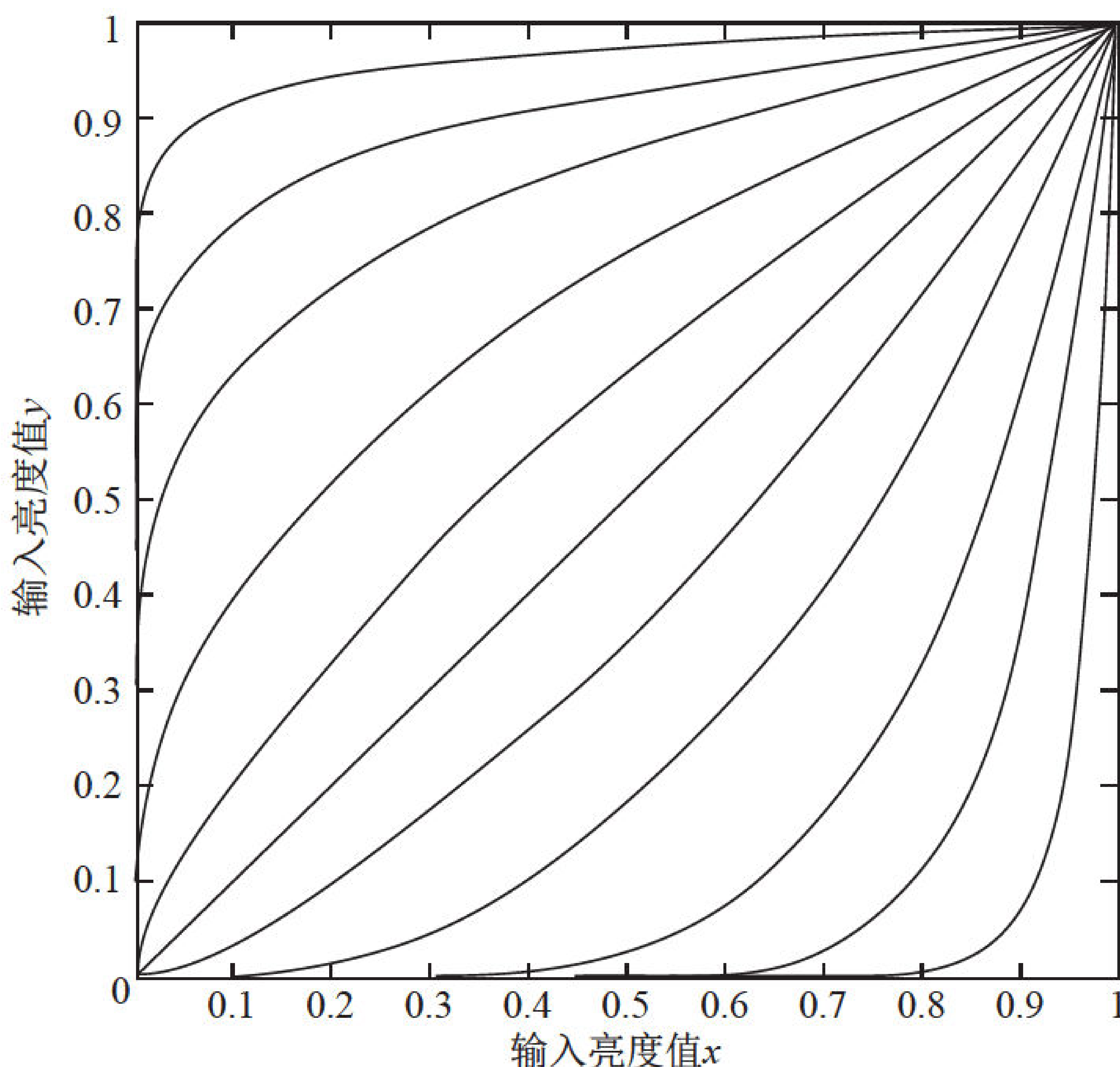


图5-1　幂次变换函数曲线

5.4.2 亮度变换函数

在两幅图像之间的亮度存在差异的情况下，可通过最小二乘法确定两幅图像亮度之间的幂次变换关系，实现伽马校正。伽马校正的数学表达形式与公式(5-20)不同，但可通过数学变换得到，以方便运算，其变换过程如下所述。

对等式(5-20)两端分别取对数，得到等式

$$\ln(y)=\ln(\beta)+\gamma\ln(x) \tag{5-21}$$

引入变量变换为

$$Y=\ln(y)，X=\ln(x)，B=\ln(\beta)，A=\gamma \tag{5-22}$$

形成线性关系式

$$Y=AX+B \tag{5-23}$$

在获得了多组(X，Y)的采样值后，即参照公式(5-20)求解A和B，则伽马校正的参数γ、β值可用下式计算

$$\gamma=A，\beta=\mathrm{e}^{B} \tag{5-24}$$

从而可以得到公式(5-20)所示的两幅图像之间的亮度变换函数关系，即伽马校正关系。

5.4.3 平滑计算

在两幅图像的重叠区域内，按从上到下、从左到右的顺序获取N个小的图像区域(大小均为$K\times L$)的亮度平均值，令在第一幅图像重叠区域中所获取的平均值为$y(i)$，i=1，2，3…N，在第二幅图像中所获取的图像平均值为$x(i)$，i=1，2，3…N。$\{(x_i，y_i)\}_{i=1}^{N}$是XY平面上的N个点的数据集合，通过引入式(5-22)的变量代换，数据线性化变换为XY平面上的点集(X_i，Y_i)。可根据式(5-20)计算相应的A值和B值，根据式(5-24)计算得到最终的亮度变换函数相应的γ值和β值。

在获得两幅图像之间的亮度变换函数之后，可对第二幅图像的亮度进行变换，使亮度数值与第一幅图像的亮度数值保持一致。具体调整时需要对像素颜色的各个分量数值分别进行调整。设变换前的像素颜色值为(r，g，b)，变换后的像素颜色值为(r'，g'，b')，则

$$(r'，g'，b')=(\beta r^{\gamma}，\beta g^{\gamma}，\beta b^{\gamma}) \tag{5-25}$$

实际应用中，可按比例调整RGB各颜色分量数值，也可考虑针对各个颜色通道建立变换函数并分别调整。

5.5　实验结果与分析

对航空拍摄得到的两幅彩色图像进行拼接计算，图像大小为600×800，如图5-2、图5-3所示。从视觉上可以明显看出两幅图像之间在亮度上存在差异。采用本书5.3节中所提出的特征模板选择方法，获得特征模板图像，如图5-4所示。采用本书提出的图像匹配计算方法，搜索得到的匹配图像，如图5-5所示。匹配后不进行亮度平滑计算得到的拼接图像，如图5-6所示。采用小波金字塔算法的图像平滑结果，如图5-7所示。采用本书所提出的方法进行全局亮度计算，得到平滑后的图像，如图5-8所示。平滑计算中，在两幅图像重叠区域采样得到12组亮度区域均值数据，如表5-1所示。通过最小二乘法计算得到的伽马校正参数分别为：γ=0.356，β=0.95。



图5-2　第1幅图像



图5-3　第2幅图像



图5-4　特征模板图像



图5-5　匹配搜索到的图像



图5-6　图像拼接结果

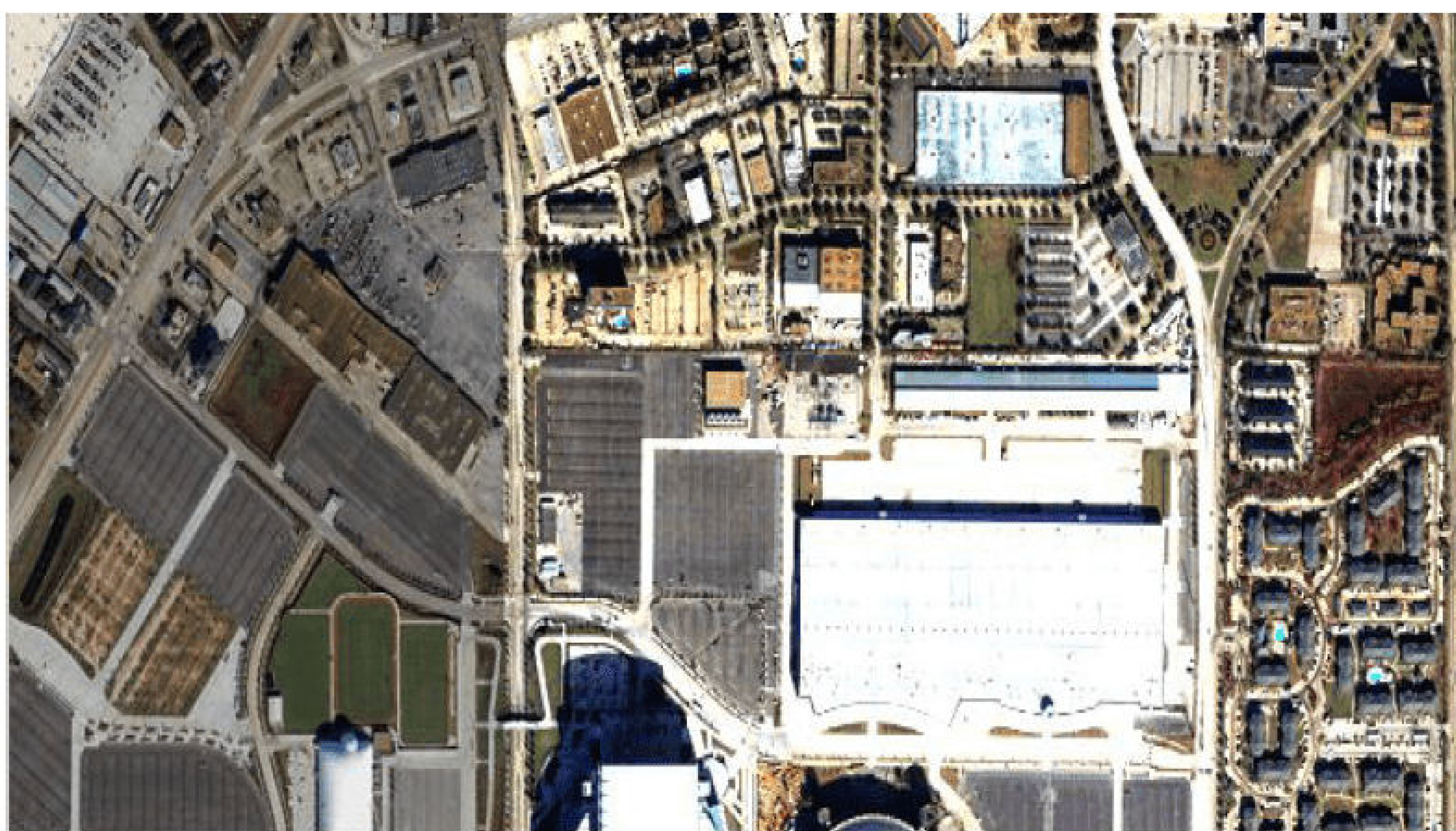


图5-7　基于小波金字塔算法的图像平滑结果

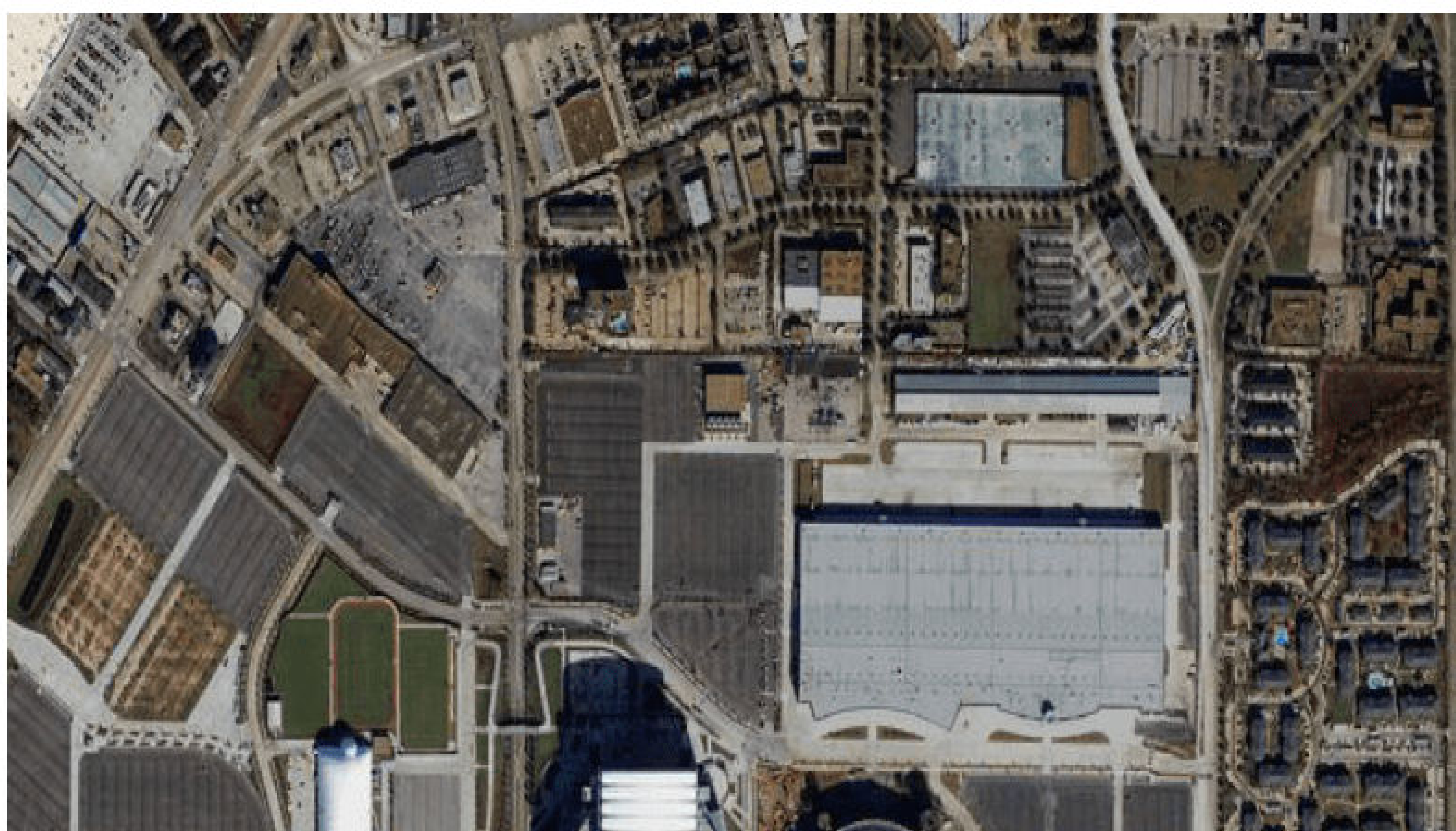


图5-8　采用本书所提出方法的图像拼接

表5-1　重叠图像区域内采样的亮度均值

序号	1	2	3	4	5	6	7	8	9	10	11	12
第1幅	115.55	56.81	79.77	81.37	86.08	81.33	78.98	89.58	130.53	87.91	94.85	158.36
第2幅	165.42	63.29	99.78	103.99	111.97	113.49	92.16	113.65	202.43	113.98	122.36	249.10

针对野外，特别是水域、大面积种植区等特征不明显的区域，使用本书提出的基于匹配以及图像平滑的方法，也能得到较好的拼接效果。

连续5幅序列图像、直接拼接后图像、使用本书方法拼接后的图像，如图5-9、图5-10、图5-11所示。

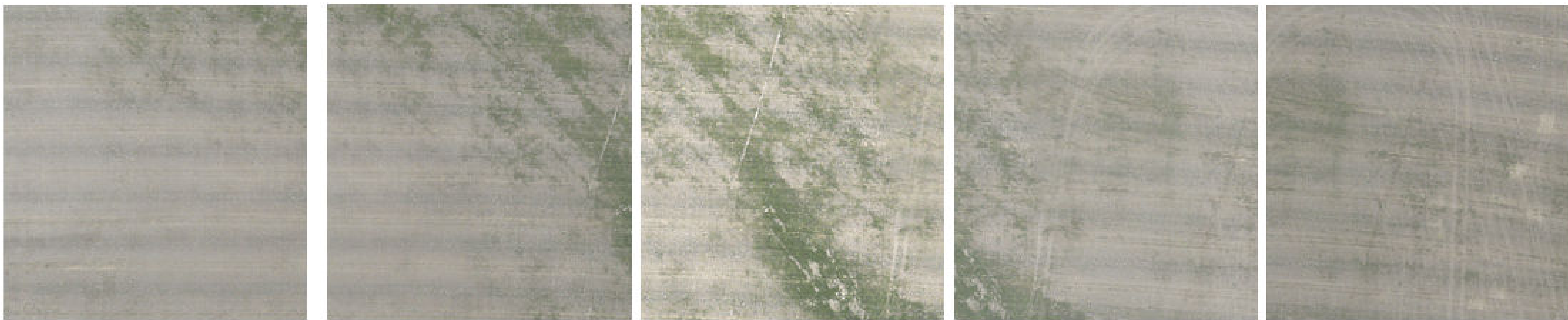


图5-9　连续5幅序列图像(600×800)



图5-10　直接拼接后图像



图5-11　使用本书方法拼接后的图像

5.6 结论

本章基于对RGB与HSI颜色空间的特性分析，提出一种根据颜色的色度和饱和度特征来度量像素颜色相似性的方法，同时提出了一种用于彩色图像匹配的特征模板提取方法。通过待拼接的两幅彩色图像之间的重叠区域亮度关系分析，建立了图像之间的亮度变化函数，全局化地调整两幅图像之间的颜色差异，消除拼接缝，增强视觉效果。实验结果表明，本书所提出的方法算法简单，计算效率以及鲁棒性均较好，对亮度存在差异的序列图像具有较好的拼接效果。

第6章

图像特征提取



特征提取是目标跟踪中非常重要的一个环节，伴随着分类器的训练和识别的全过程，在分类器的训练中起到至关重要的作用，直接影响最终的跟踪效果及整个系统的性能。目标信息获取越多，则目标外观表达越准确，跟踪效果越好，但同时计算复杂度就越高，所需的处理时间也相应越长，使算法的实时性得不到保证。研究者希望选择折中目标描述信息中易于区分且数据量小的特征信息，这个过程就是特征提取。

6.1 图像特征

由于在基于分类学习的目标跟踪算法中，训练分类器的过程复杂，所以对于目标特征的选择就有更高的要求。如果特征的数据量较大，跟踪算法无法保证过程的实时性；如果特征的数据量太小，目标信息量较少，将难以准确区分目标与非目标。所以，基于分类学习的目标跟踪一般不选择较直观的整体特征，而偏向于由基本特征进行局部特征统计，使特征描述子更易于区分且计算量较小。这些特征有其各自的特点和应用场合。下面介绍几种目标检测中常用的图像特征。

6.1.1 LBP特征

Ojala等人[192][193]提出的局部二进制模式(Local Binary Pattern，LBP)，其显著优点在于运算简单且具有旋转不变性和灰度不变性，被广泛应用到图像分类与检索等问题中。[34]目前，LBP算子能有效描述图像纹理，在指纹识别、人脸识别等领域都已得到成功应用。LBP主要是针对灰度图像而言，它统计一幅图像中每一个像素与周围邻域内像素的对比关系，因此LBP对光照变化不敏感，但容易受到噪声的影响。

LBP算子的最初定义是在3×3的窗口中，将中心单元格像素值设定为阈值，并将该像素值与周围相邻的8个像素值相比较，将周围像素值相对阈值大的点设为1，反之为0。经过一轮比较，按顺时针获取中心像素8邻域生成的8位二进制数，将二进制数转换为十进制数值，该值就是窗口中心点的LBP值，如图6-1所示。

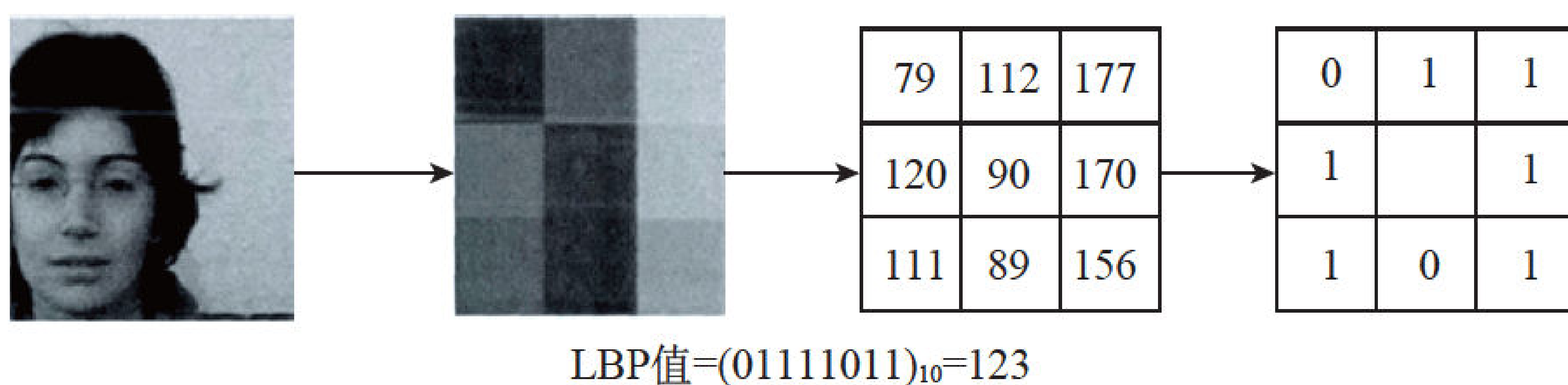


图6-1　LBP特征

LBP算子被提出来后，研究者对该算子做了进一步的优化与改进，提出了圆形LBP算子、LBP旋转不变模式、LBP等价模式。

1. 圆形LBP算子

最初LBP算子的适用范围仅是一个具有固定半径的小区域，显然，该算子不能满足不同尺寸、不同频率的纹理的需求。为了克服这一缺点，Ojala等对LBP算子做了进一步研究，将原始窗口扩大到一个以R为半径、周围有P个邻域点均匀分布的更大的区域上。该LBP算子用$\mathrm{LBP}_{P,\ R}$表示，常见的LBP算子有$\mathrm{LBP}_{8.1}$、$\mathrm{LBP}_{8.2}$、$\mathrm{LBP}_{8.3}$，如图6-2所示。

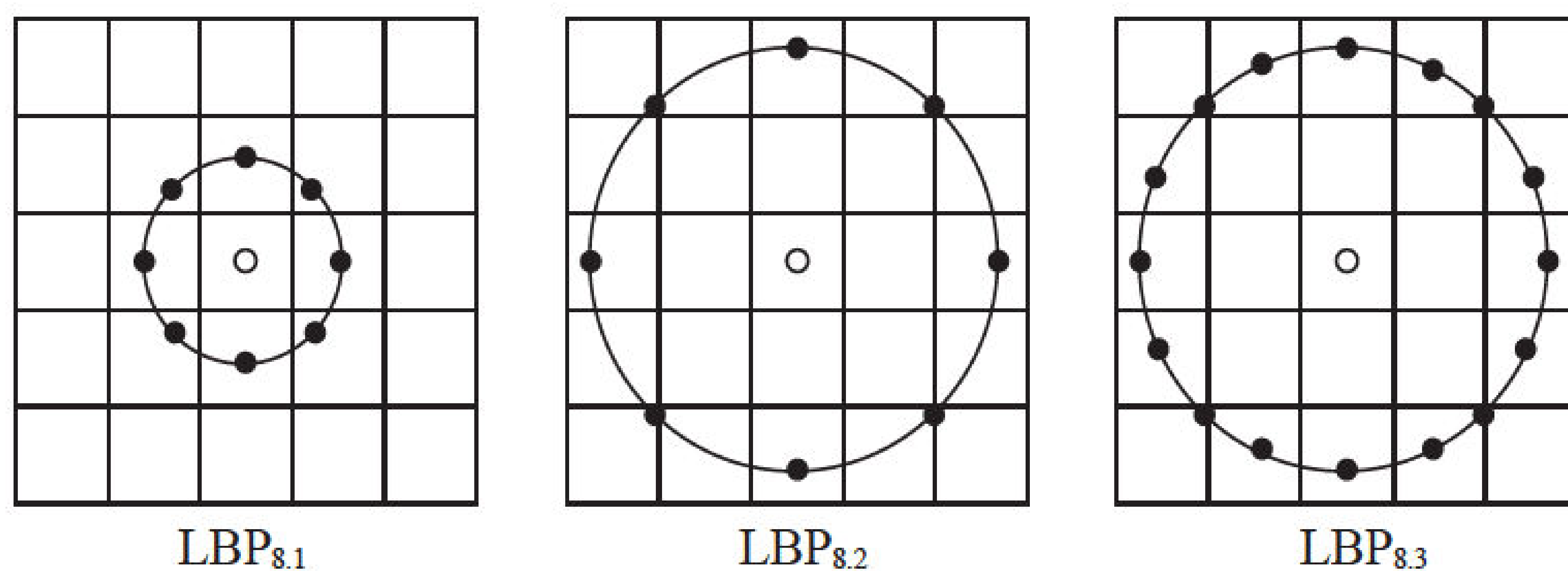


图6-2　常见的LBP算子

假设，g_c为中心点像素值，以该点为圆心、$R(R>0)$为半径的圆上，均匀分布$P(P>1)$个相邻点，像素值分别设为g_0，g_1，…，g_{p-1}。LBP算子会将g_c与圆周上的P个点的像素值相比较，按顺时针得到8位二进制串，最后将二进制串转为十进制数值，该数值就是中心像素点的LBP值。

$$\mathrm{LBP}_{P,R}=\sum_{P=0}^{P}s\left(g_P-g_c\right)2^P \tag{6-1}$$

这里

$$s(x)=\begin{cases}1 & x\geqslant 0\\0 & x<0\end{cases} \tag{6-2}$$

同理，对整幅图像的每个像素点，应用上述方法都可得到相应的LBP值。最后利用直方图对整幅图的所有像素点的LBP值进行统计并得出统计向量，该向量就是整幅图像的LBP特征。我们可以看到LBP特征是基于像素差值的统计，所以LBP特征对于灰度单调变化的图像具有很好的鲁棒性，同时具有平移不变性。[194]

2. LBP旋转不变模式

从以上理论很容易得知，LBP算子的像素值是固定值，具有灰度不变性，但一旦图像稍微旋转，LBP值就会发生改变，所以LBP算子不具有旋转不变性。为了克服这个问题，研究者将LBP算子进行了改进，他们对圆形邻域做多次旋转，相应地会得到多个LBP值，将最小值作为该邻域的LBP值。采用该方法得到的局部二值模式就是旋转不变LBP算子。

图6-3描述了旋转不变LBP值的获取过程。每个算子图下面都有其LBP值与之对应。图6-3给出的8种模式都是经过不同的旋转得到的，我们最终选取最小值15作为旋转不变性的LBP值，其对应模式为00001111。

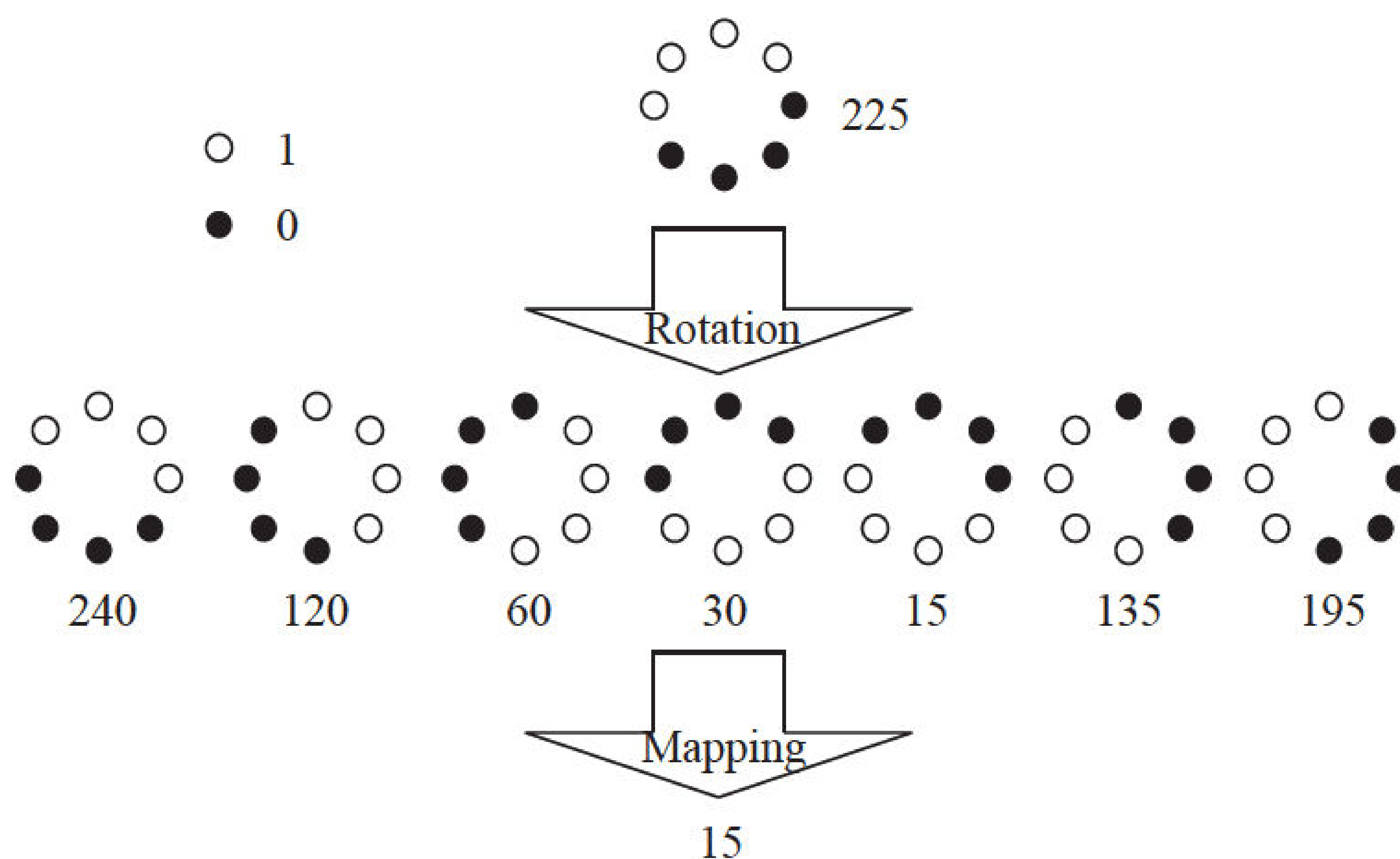


图6-3　旋转不变的LBP示意图

3. LBP等价模式

对于以R为半径且具有P个采样点的圆形区域，会有2^P种二值模式与LBP算子相对应。如果P值增加，LBP算子所对应的二值模式种类也会增加。对这么多模式进行纹理提取与分类是非常困难的，而且过多的冗余信息也会对纹理的表达有很大的影响。因此，解决此问题的最好办法是在保证图像特征信息的同时，通过降维来减少LBP算子的模式种类。

研究者为了减少数据量，对LBP模式进行改进优化。在实际应用中，LBP模式通常最多只有两次0，1间的跳变。因此，定义“等价模式”：如果某个LBP对应的循环二进制数

最多只有两次0，1间的跳变，则该二进制模式就被称为等价模式类。除等价模式类以外的其他模式都统一称为混合模式类。我们用U表示P个采样点生成的二进制串中0/1和1/0跳变的次数。

$$U(\mathrm{LSB}_{P,\ R})=\left|s(g_{p-1}-g_c)-s(g_0-g_c)\right|+\sum_{p=1}^{p-1}\left|s(g_p-g_c)-s(g_{p-1}-g_c)\right| \tag{6-3}$$

我们只需要统计跳变次数小于或等于2情形下的U值所对应的LBP值，其他都属于另一类。

$$\mathrm{LBP}_{P,\ R}^{\mathrm{uniform}}=\begin{cases}\sum_{p=0}^{p-1}s(g_p-g_c)2^p & U(\mathrm{LSB}_{P,\ R})\leqslant 2\\ x & \text{otherwise}\end{cases} \tag{6-4}$$

通过以上方法，我们可以大大降低二值模式的种类，并且对信息量无任何影响。对于P个采样点的情况，二值模式的数量由原来的2^P降为$P(P$□$1)+2$。

另外，将旋转不变LBP与等价LBP结合起来便得到旋转不变等价LBP模式(Rotation Invariance Uniform LBP)，其含有的模式个数为$P+1$，如图6-4所示。该方法达到了对特征向量降维的目的，同时也可有效抑制高频噪声。

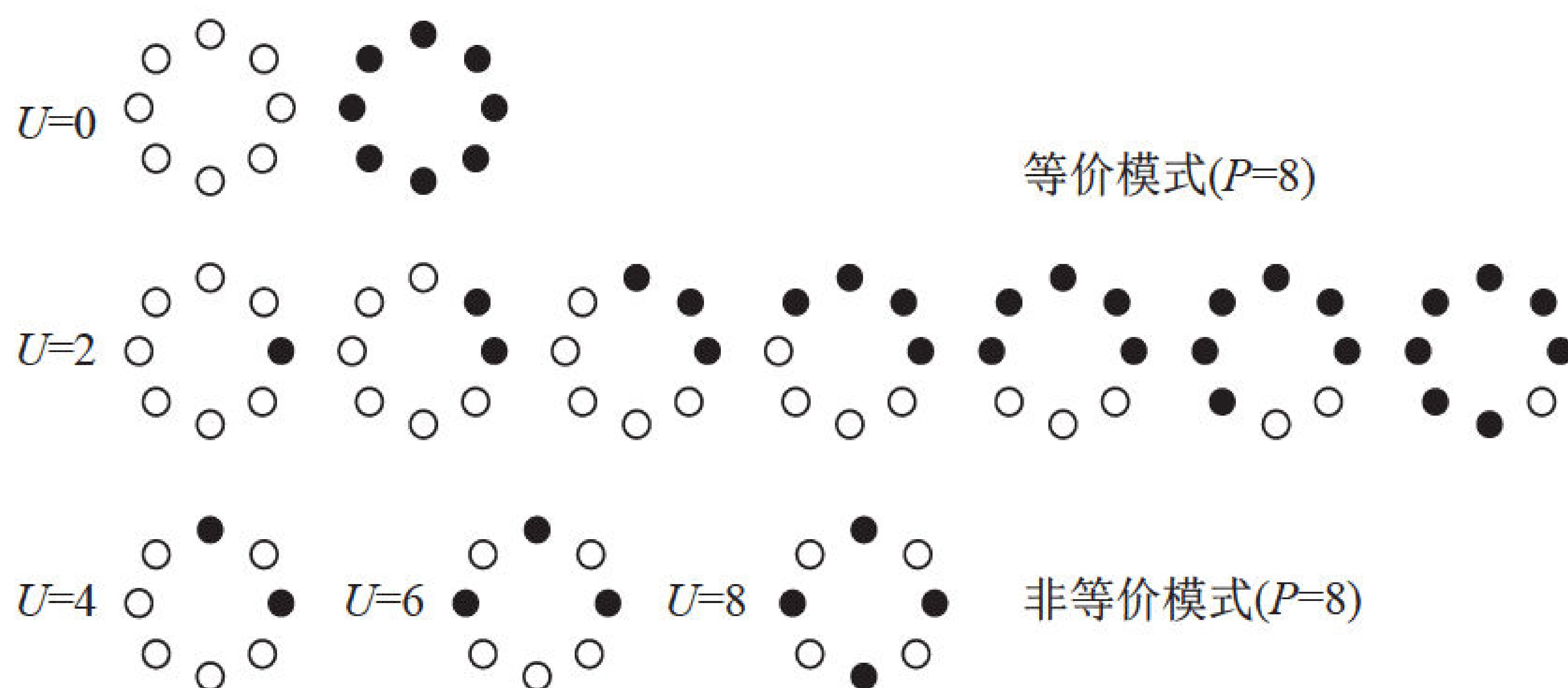


图6-4 LBP算子等价模式和非等价模式

6.1.2 Haar-like特征

在计算机视觉领域中，经常用Haar-like特征算子来描述目标的特征。它最早是由Papageorgiou[195]等人提出用于人脸描述，后来扩展到目标识别领域的各种方法中。但Haar-like特征的最大缺点是计算复杂。

Haar-like特征主要有线性特征、边缘特征、点特征(中心特征)三种常用的类型，如

图6-5所示。

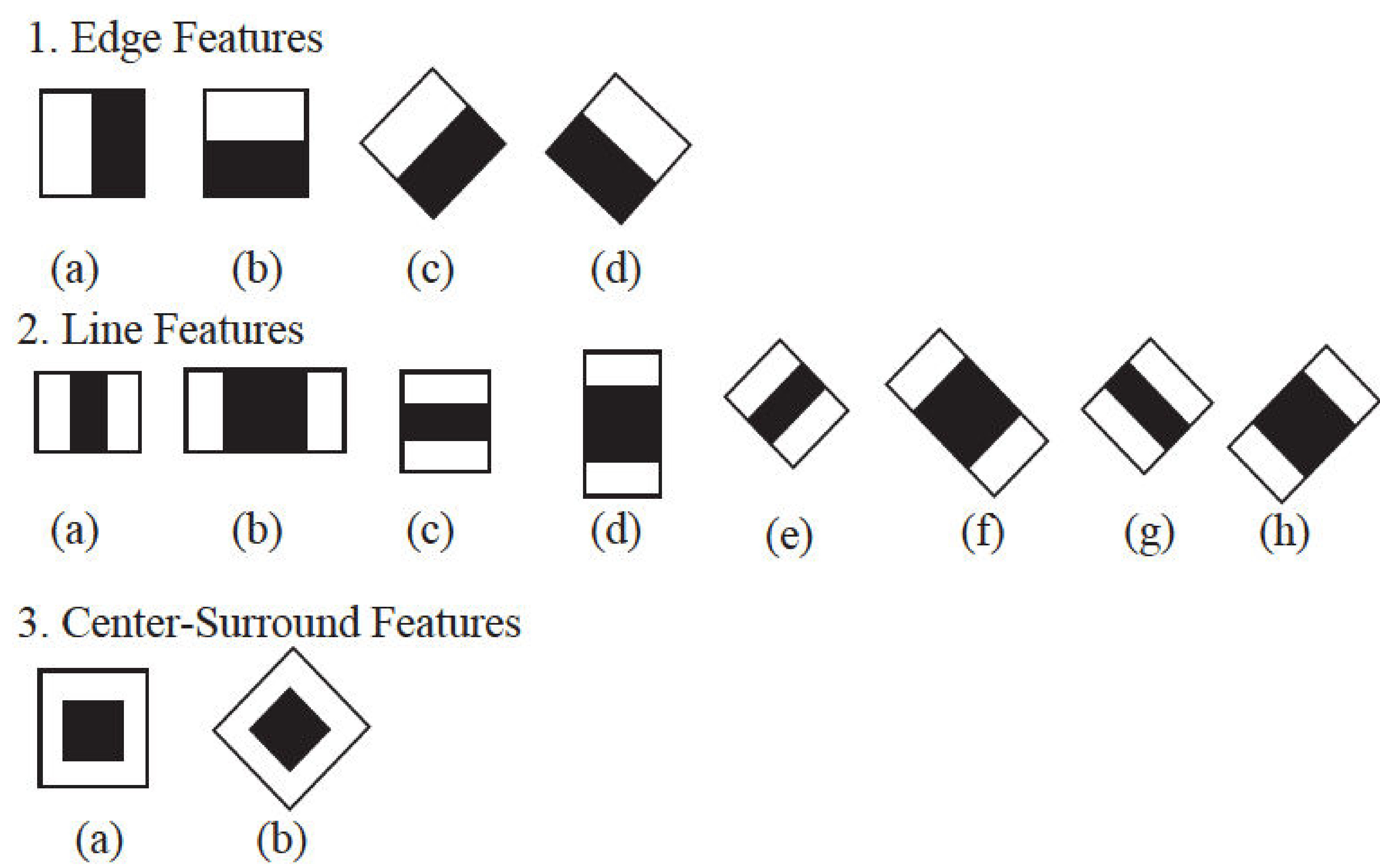


图6-5　三种Haar-like特征类型

显然，它包含4种边缘特征，即x方向、y方向、x倾斜方向、y倾斜方向；8种线特征；两种点特征。但在使用Haar-like特征时，我们一般常选a、b、c、d4种类型，它的优点在于计算量相对较小，如图6-6所示。

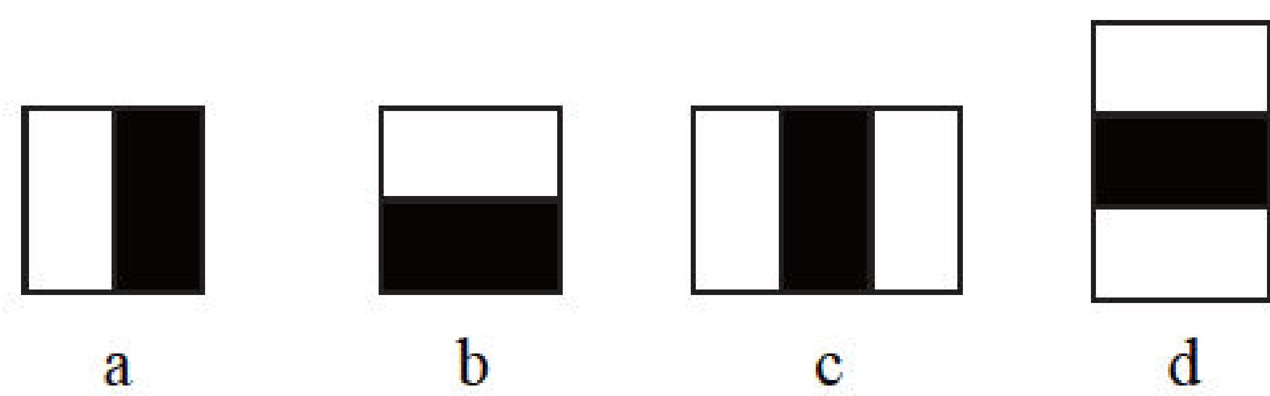


图6-6　4种常用的Haar-like特征

计算特征值时，黑色区的权值设为负，白色区的权值设为正。因此，a、b类型的特征值计算公式为

$$V=\mathrm{Sum}_{白}-\mathrm{Sum}_{黑} \tag{6-5}$$

后两种类型的特征值计算公式为

$$V=\mathrm{Sum}_{白}-2*\mathrm{Sum}_{黑} \tag{6-6}$$

式(6-6)中黑色区域像素乘以2，其目的在于使黑白两个区域的像素个数保持一致。[196]

Haar-like特征用于检测的原理如下：

将上述方法应用到人脸检测中，我们在上述4种矩形中任选一种与人脸区域相匹配，求白色区与黑色区的像素和之差，该差值就是人脸特征值。把该矩形用来与非人脸区域相匹配，并求其特征值，那么该值与人脸特征值会有很大的差别。我们希望相差越大越理

想。在上述求解过程中，矩形块把人脸特征进行量化，根据特征值来确定该区域是否为人脸，并进行人脸识别。

6.1.3　SIFT特征

尺度不变特征变换(Scale Invariant Feature Transformation，SIFT)特征，是由Lowe[197]提出的一种提取局部特征的算法。SIFT特征是在多尺度空间分析下进行特征检测，然后形成特征描述符，以达到对尺度缩放、旋转变换、放射变换，具有不变性。如今，该算法已成功应用于图像匹配和图像拼接领域中。

下面对SIFT算法的4个主要步骤进行介绍。

1. 尺度空间极值点检测

设一幅二维图像$I(x，y)$，它与高斯核卷积可以得到不同尺度下的尺度空间

$$L(x, y, \sigma)=G(x, y, \sigma)*I(x, y) \tag{6-7}$$

其中，$G(x，y，\sigma)$是高斯函数，定义为$G(x, y, \sigma)=\frac{1}{2\pi\sigma^2}e^{-(x^2+y^2)/2\sigma^2}$。

为了得到稳定特征点并使特征点具有尺度不变性，我们需要在图像的二维平面空间和高斯差分(Difference of Gaussian，DoG)尺度空间中同时检测局部的极值点。于是，DoG算子的定义为

$$\begin{aligned}D(x, y, \sigma)&=[G(x, y, k\sigma)-G(x, y, \sigma)]*I(x, y)\\&=L(x, y, k\sigma)-L(x, y, \sigma)\end{aligned} \tag{6-8}$$

在DoG尺度空间中检测极值的时候，对DoG图像中的点与其邻域的相邻点相比较，包括自身8邻域和相邻上下尺度层图像对应的9邻域相比较，如图6-7所示。若该点为极值，则将其选为特征点。

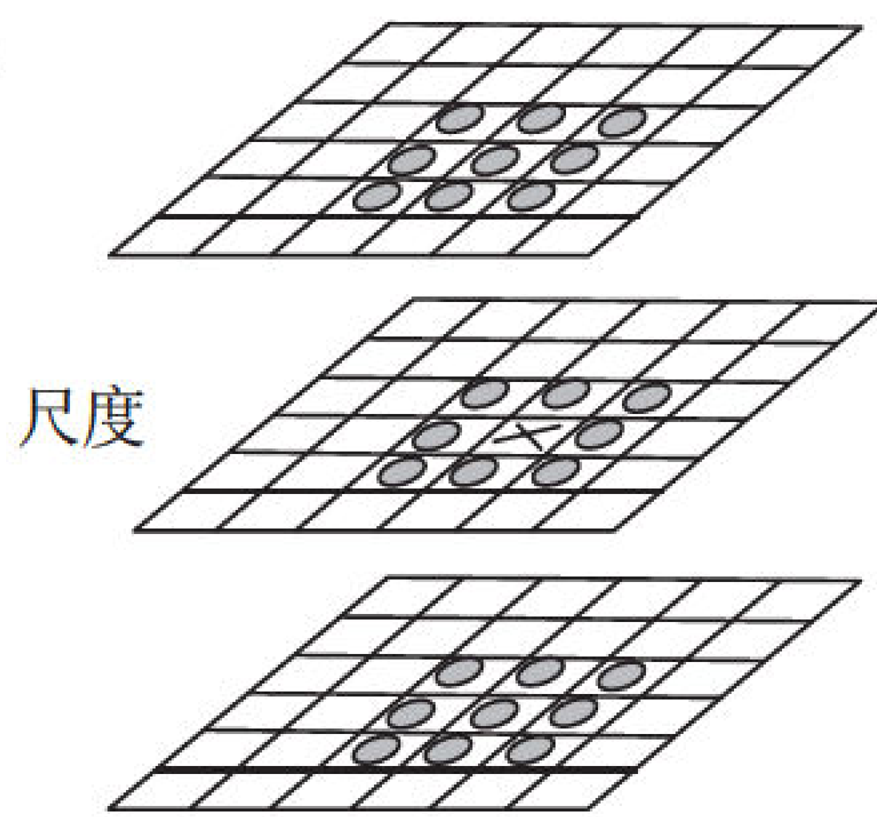


图6-7　DoG尺度空间极值点检测

2. 特征点的定位

由于DoG算子会产生较强的边缘响应，因此特征点定位就是要去除极值点中对噪声敏感的点或者不稳定的边缘响应点。我们采用具体方法对空间尺度函数$D(x，y，\sigma)$进行泰勒展开，可得到精确的特征点位置和尺度，并且还可以滤除低对比度的特征点。高斯差分函数$D(x，y，\sigma)$在局部极值点$D(x_0，y_0，\sigma)$处的泰勒展开式为

$$D(x,\ y,\ \sigma)=D(x_0,\ y_0,\ \sigma)+\frac{\partial D^T}{\partial x_0}+\frac{1}{2}x^T\frac{\partial^2 D}{\partial x^2}x \tag{6-9}$$

对公式(6-9)求导并令其为零，最终求得的x值就是精确位置，如公式(6-10)所示。

$$x_{\max}=-\left(\frac{\partial^2 D}{\partial x_0^2}\right)^{-1}\frac{\partial D}{\partial x_0} \tag{6-10}$$

设对比度的阈值为0.03，如果$|D(x_{\max})|\geqslant 0.03$则该特征点给予保留，否则认为该极值点为低对比度点，给予丢弃。

对于一个不理想的高斯差分算子，极值处的曲率随方向的不同变化很大，如横跨边缘的方向值很大，在垂直边缘的方向值却非常小。我们可以通过2×2的Hession矩阵计算主曲率值来去掉不稳定的边缘点，即

$$H=\begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \tag{6-11}$$

因为D的主曲率和H的特征值成正比，所以可依据Hession矩阵的秩$T_r(H)$和行列式$\mathrm{Det}(H)$的比值，判断极值点是否为不稳定的边缘点。设最大的特征值为λ，最小的特征值为γ，则

$$\begin{aligned} T_r(H)&=D_{xx}+D_{yy}=\lambda+\gamma \\ \mathrm{Det}(H)&=D_{xx}D_{yy}-(D_{xy})^2=\lambda\gamma \end{aligned} \tag{6-12}$$

令$\lambda=\mu\gamma$，即μ为λ和γ的比值，则

$$\frac{T_r(H)^2}{\mathrm{Det}(H)}=\frac{(\lambda+\gamma)^2}{\lambda\gamma}=\frac{(\mu\lambda+\gamma)^2}{\lambda\gamma^2}=\frac{(\mu+1)^2}{\mu} \tag{6-13}$$

当$\lambda=r$时，$(\mu+1)^2/\mu$的值最小，并且当μ增大时，$(\mu+1)^2/\mu$也会随之增大。因此，我们只需检测$\frac{T_r(H)^2}{\mathrm{Det}(H)}<\frac{(\mu+1)^2}{\mu}$是否成立，就可得知主曲率与阈值$\mu$的大小关系。

3. 特征点方向的确定

特征点邻域像素的梯度方向分布决定了该点的方向，因此，SIFT算子对图像的旋转具有不变性。特征点(x, y)处梯度的大小和方向的计算公式为

$$d(x, y)=\sqrt{[L(x+1, y)-L(x-1, y)]^2+[L(x, y+1)-L(x, y-1)]^2} \tag{6-14}$$

$$\theta(x, y)=\arctan[L(x, y+1)-L(x, y-1)]/[L(x+1, y)-L(x-1, y)] \tag{6-15}$$

L代表每个特征点所在的尺度。

4. SIFT特征向量的生成

首先，旋转坐标轴方向，使之与特征点方向一致，确保图像具有旋转不变性。其次，以特征点为中心取164×416个网格窗口，并分割成16个4×4子网格区域，每个子网格都被称为一个种子点。对于每个子网格区域，我们都需要计算8个梯度方向和梯度模值，并进行梯度直方图统计和归一化处理。一个子块总共可以得到164×48=128个方向的描述符，这个长度为128的一维向量被定义为SIFT特征向量。最后，对得到的最终特征向量进行归一化操作，就可以获取SIFT特征描述子，具体的形成过程如图6-8所示。

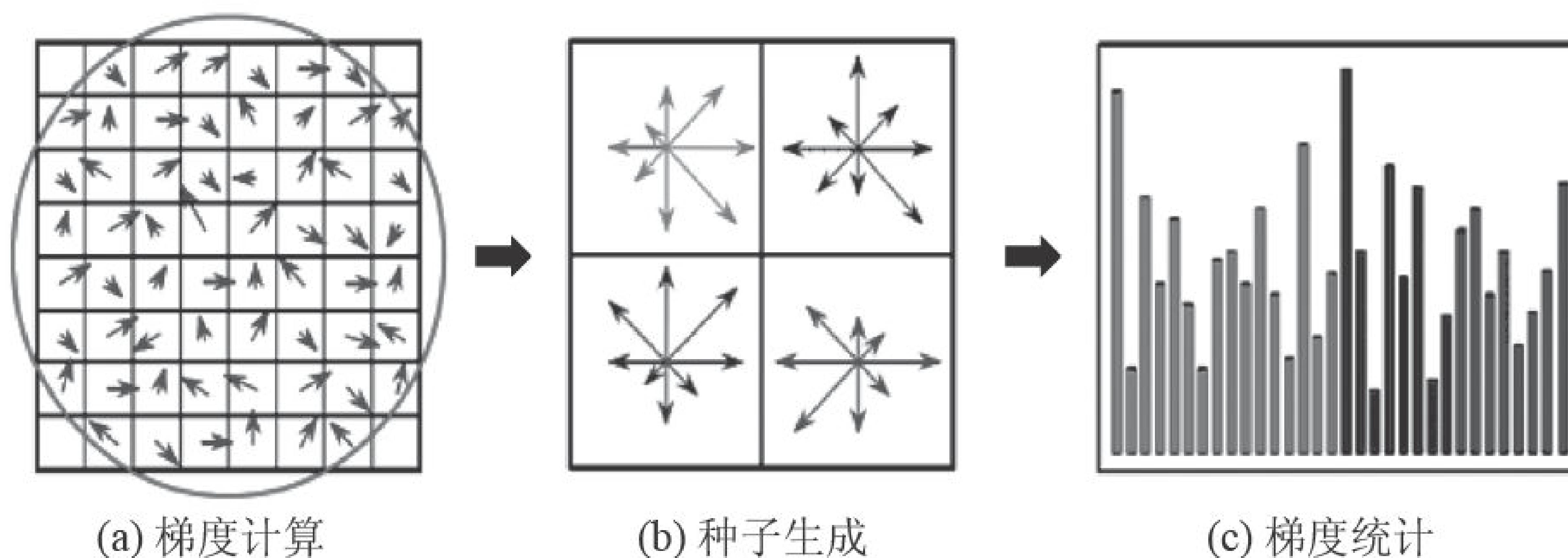


图6-8　SIFT特征形成过程

6.1.4　HOG特征

方向梯度直方图(Histogram of Oriented Gradient，HOG)是由N.Dalal和B.Triggs[198]等人于2005年提出的一种基于梯度方向统计的特征，最初HOG被用于静态图像或者视频中的行人检测，随后被广泛应用于各种对象，包括刚体如汽车、摩托车和非刚体如动物和人的检测、跟踪和识别。

HOG是用于目标检测的一种特征描述子，广泛应用于计算机视觉和图像处理领域。HOG特征的主要思想是通过提取局部梯度方向的分布，有效地描述图像中目标的局部外

观形状，而无须寻找确定边缘的位置。在特征提取中，检测图像窗口被划分为若干个单元格(Cell)，在每个单元格内统计各像素点一维梯度方向的直方图。[199]最后，基于更大的区域块进行归一化处理，形成图像目标的特征表示。

6.2 HOG特征的提取方法

HOG特征提取流程，如图6-9所示。

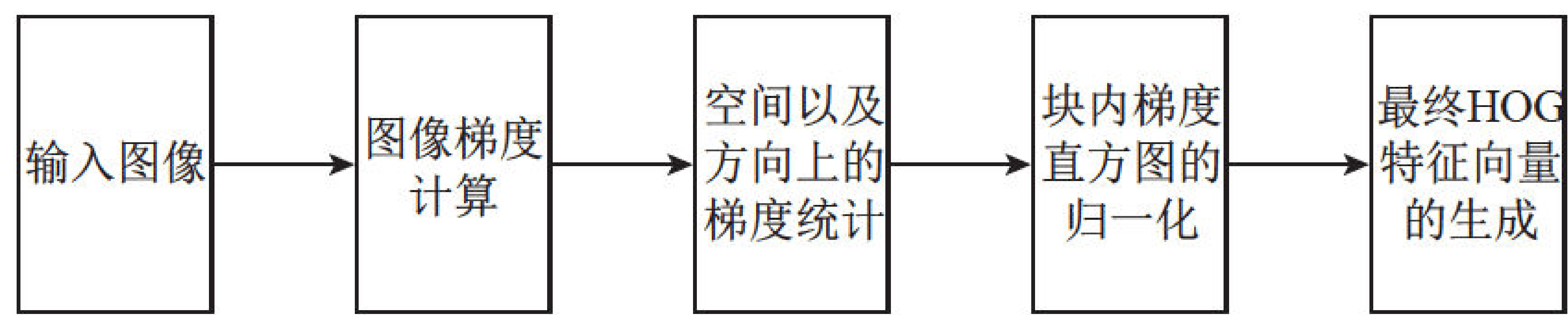


图6-9 HOG特征提取的基本流程图

6.2.1 图像梯度计算

图像中的每个像素都可以用该像素的梯度方向和梯度幅值表示，如图6-10所示。

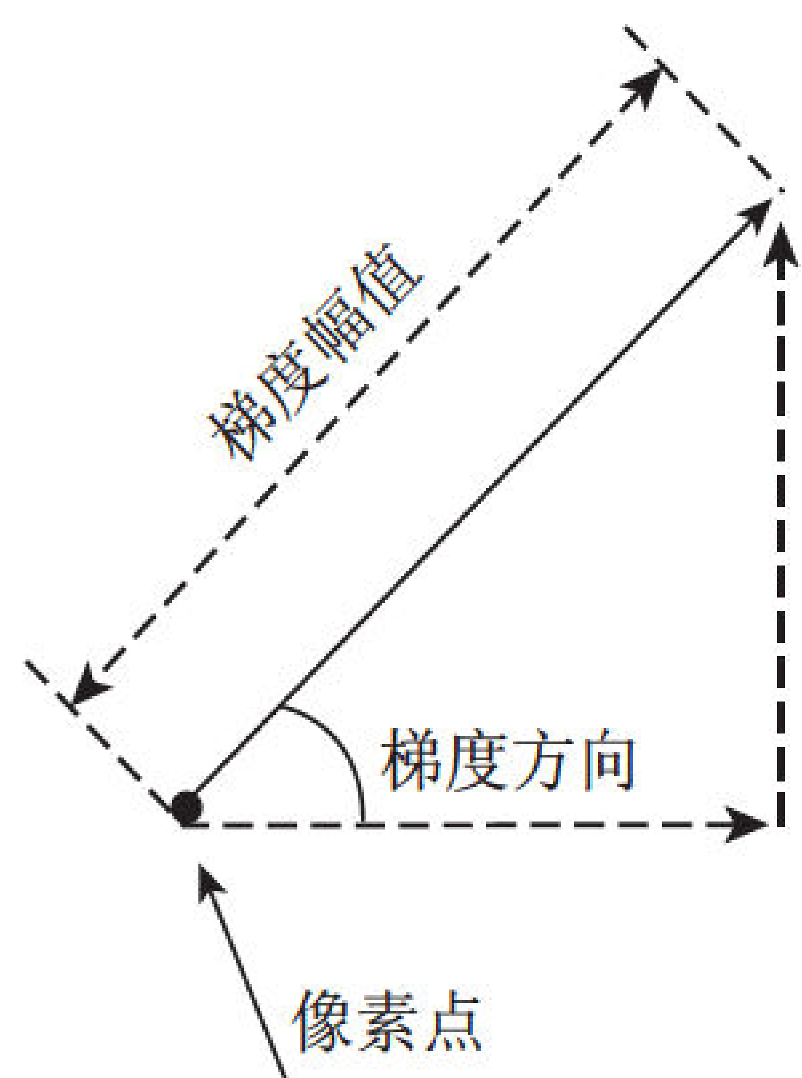


图6-10 像素点的梯度幅值和方向

一般情况下，像素的梯度是根据该像素上横坐标方向与纵坐标方向的梯度计算求得。从Dalal和Triggs的测试结果可知，使用一维的离散梯度模板[-1 0 1]实现起来比较简单、快

捷、有效，可以方便地获得图像梯度信息。而其他较复杂的离散的微分模板，如Prewitt或者Sobel，在其运算量增加的同时，得到的检测效果不是很理想。

假设一幅图像I在像素点$(x，y)$处的像素值用$I(x，y)$表示，则$(x，y)$处的水平和竖直梯度计算公式为

$$G_x(x，y)=I(x+1，y)-I(x-1，y) \tag{6-16}$$

$$G_y(x，y)=I(x，y+1)-I(x，y-1) \tag{6-17}$$

样本点$(x，y)$的梯度幅值为

$$G(x，y)=\sqrt{G_x(x，y)^2+G_y(x，y)^2} \tag{6-18}$$

样本点$(x，y)$的梯度方向为

$$\theta(x，y)=\tan^{-1}\left(\frac{G_y(x，y)}{G_x(x，y)}\right) \tag{6-19}$$

6.2.2　空间以及方向上的梯度统计

对每个单元格内像素点的梯度量化进行直方图统计。梯度方向的角度范围可以为0至180°(无符号梯度——梯度向量的正负值被映射到同一个方向组)，也可以是0至360°(有符号梯度)。实验结果表明，梯度方向的角度范围是[0，180°)，并且直方图通道(Bin)为9时表现得最好。根据梯度方向的范围将梯度直方图划分为9个部分(每个部分对应一个梯度方向区域，每20°为一个区域；计算像素点的幅值与方向，在像素点落入方向的区域中进行梯度幅值叠加)，如图6-11所示。

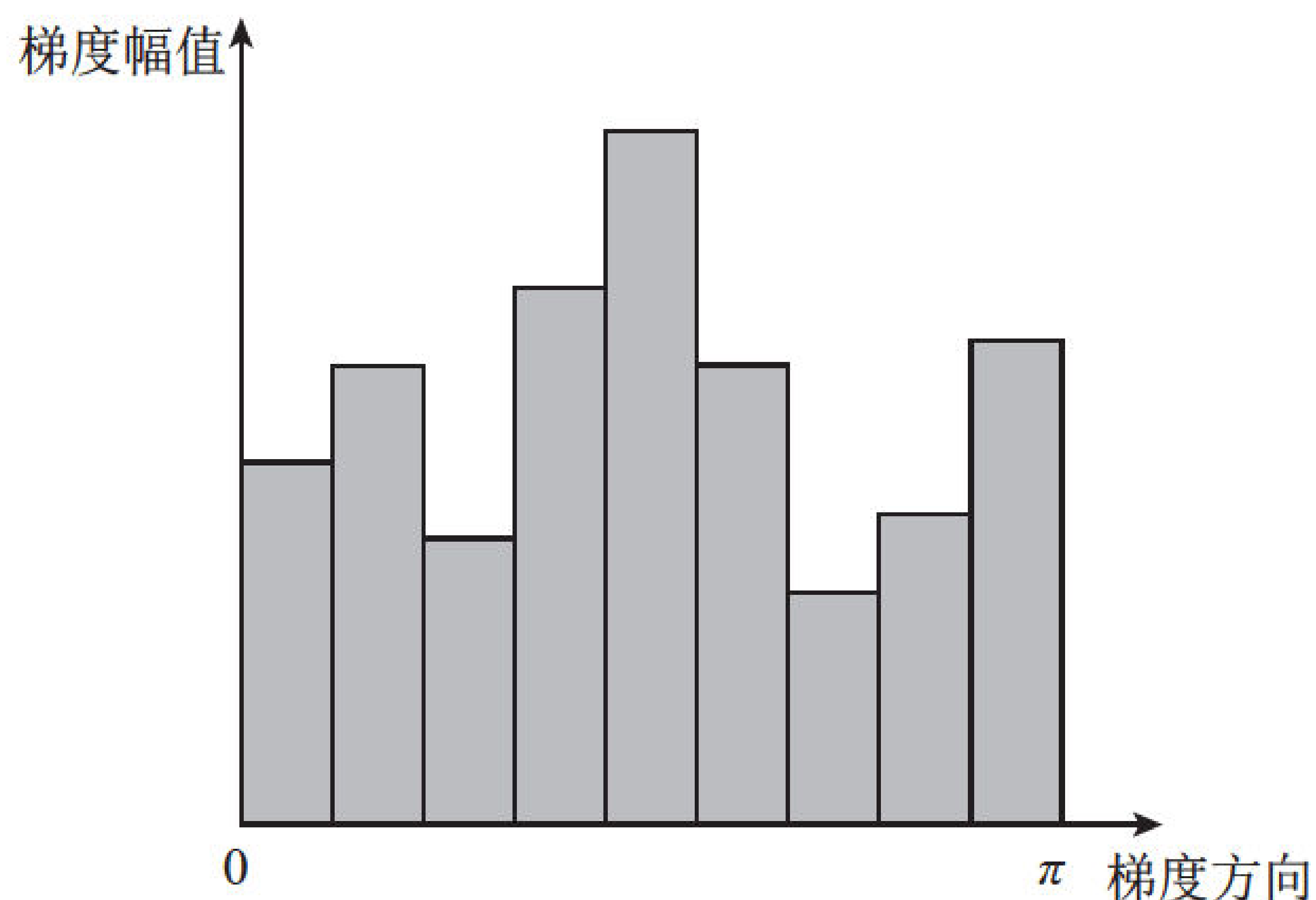


图6-11　梯度直方图划分的9个部分

统计梯度方向直方图是以各单元格(Cell)进行的。一般情况下，每个细胞单元(Cell)按照8×8像素来划分，然后再将每相邻的4个小细胞单元组成一个小的区域，我们称之为Block，每个块(Block)为2×2个Cells。常用的块(Block)的形状可以是矩形(R-HOG)或者圆形(C-HOG)，本书采用的是矩形，如图6-12所示。它可以用三个参数来表征：块(Block)中细胞单元的个数、细胞单元(Cell)中像素点的个数、细胞单元中直方图通道(Bin)的个数。

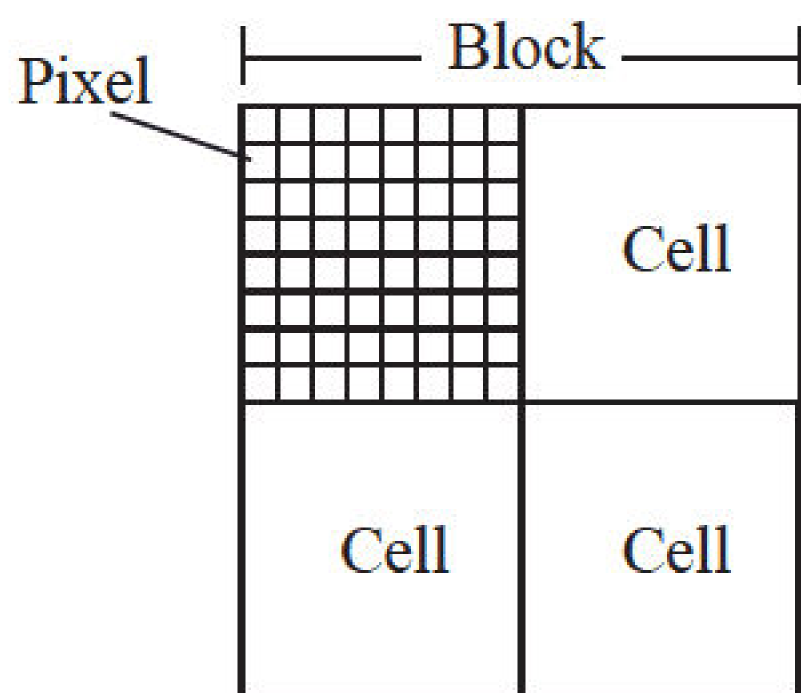


图6-12 Block与Cell的示意图

要计算每个细胞单元内对应的梯度方向直方图，需要采用加权统计的方式来进行，因为考虑到每一个像素的方向梯度不一定都能落在区域中央。假设某像素点的梯度角度落在区域[20°，40°)边缘附近20.2°的方向，这时候不能忽视这个像素点梯度幅值对于临近区域[0，40°)的影响。所以，如果不对梯度幅值进行相邻区域的投影，将会造成特征的不稳定。这时候可以利用投影的方法，将每个像素点的方向投影到相邻的两个区域中心(10°和30°)，在相邻的区间内都进行梯度的叠加。对于加权统计的计算，其权值可通过像素点梯度的幅值或者幅值的函数来表示，幅值的函数可以采用幅值的平方根或者平方等。根据Dalal等的实验结果可知，采用梯度幅值作为权值，进行加权统计可获得最好的效果。为了避免梯度方向直方图在Cell单元格边界以及Bin区间边界处发生突变，在相邻方向组采用三线性插值，即把当前像素的梯度幅值作为插入值，把像素的梯度方向、在Cell单元格中的x坐标与y坐标这三个值作为插值权重进行权值分配。

这样，每个Cell单元格首先通过对各像素点的梯度幅值加权投影构成9维的特征向量，再对单元格内各梯度方向进行投影操作，即映射到直方图通道中相应的角度范围，该Cell单元格的HOG特征描述子就已形成。

6.2.3 块内梯度直方图的归一化

图像序列不同的图片之间或同一图片的不同区域之间的梯度方向直方图有时会相差很

大，这主要是因为局部光照强度的不同以及前景与后景对比度的不同。因此，为了提高检测的效果，削弱光照及前后景对比度不同的影响，我们有必要对图像进行局部对比度的归一化。

进行归一化的处理办法有很多，常用的就是把小的单元格(Cell)区域统计到一个稍大的区域块(Block)，然后基于每个块进行归一化。这里，归一化是对每个块(Block)进行的，归一化函数主要有以下4种：

(1) L1-norm：$v \leftarrow v / \sqrt{\|v\|_1 + \varepsilon}$；L1-norm就是模值相加，即曼哈顿距离。

(2) L1-sqrt：$v \leftarrow \sqrt{v / (\|v\|_1 + \varepsilon)}$；L1范数取平方根。

(3) L2-norm：$v \leftarrow v / \sqrt{\|v\|_2^2 + \varepsilon^2}$；L2-norm就是平方和开方，即欧氏距离。

(4) L2-Hys：L2范数的基础上进行限幅归一化，使v的最大值为0.2。

假定未归一化的特征描述子向量为v。$\|v\|_k$表示v的k阶范数，k=1，2。ε是一个很小的归一化常数，用来防止发生除数为0的情况。

6.2.4　最终HOG特征向量的生成

所有Block内各个Cell的特征向量连接起来构成HOG特征描述子。几个Cell会构成一个Block(如2×2个Cell对应一个Block)，每个Block生成36维的HOG特征，如图6-13所示。再对检测图像窗口滑动搜索扫描，遍历每个Block并且计算Block的特征向量，连接组合所有Block的向量就是这个窗口所对应的目标特征，也是最终可用于分类的特征描述子。

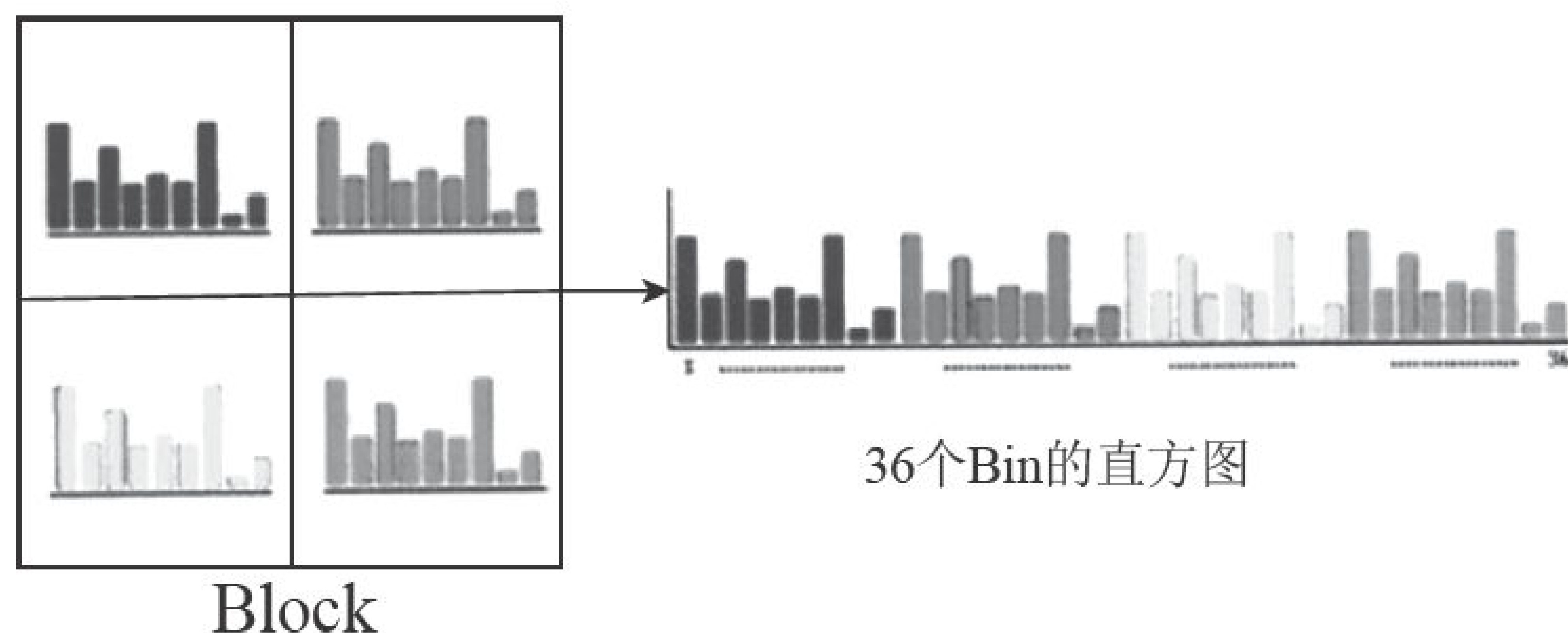


图6-13　一个Block的梯度直方图

6.3　HOG特征的目标描述

HOG特征的提取方法是按照分块(Block)分单元(Cell)进行的，所以对图像局部像素点间的关系可进行有效表达。由于HOG特征拥有几何不变性，对光照变化不敏感，HOG特征在手势识别、人脸识别、车辆检测等应用领域都表现出良好的性能，如今HOG特征被广泛应用于目标检测领域中。因此，本书提取HOG特征进行图像目标的描述表达。

图像中提取HOG特征的主要步骤，如图6-14所示。

(1) 输入图像，利用固定大小的窗口扫描图像；

(2) 提取窗口内的图像信息；

(3) 将窗口均分为多个不重叠的单元，每个单元为8×8像素；

(4) 计算每个单元内的像素横坐标方向和纵坐标方向的梯度幅值，并利用这些信息计算出该像素的梯度方向和梯度幅值；

(5) 在每个单元内按照给定的量化标准统计每个梯度方向范围内所有像素的梯度幅值，构造出直方图，从而得到这个单元的HOG特征向量；

(6) 将每个单元2×2区域内的单元称为一个块，对每个块内的单元进行对比度归一化处理，从而降低了光照的影响，每个块的HOG特征向量是由其内的每个单元HOG特征向量串联所得；

(7) 连接组合所有块内的HOG特征向量，构成图像目标的高维HOG特征描述子。

本书采用8×8个像素大小的小细胞单元Cell，4个Cell组成一个局部区域块Block，每个Cell内有9个梯度方向的维度，每个Block以8个像素宽度为步长对输入图像进行扫描，进而可以提取出每个Block内的HOG特征，每个Block内有4个9维的梯度特征向量，对块中所有直方图采用L2-Hys归一化函数实现归一化，最终可以提取出每个样本中高维的HOG特征向量。

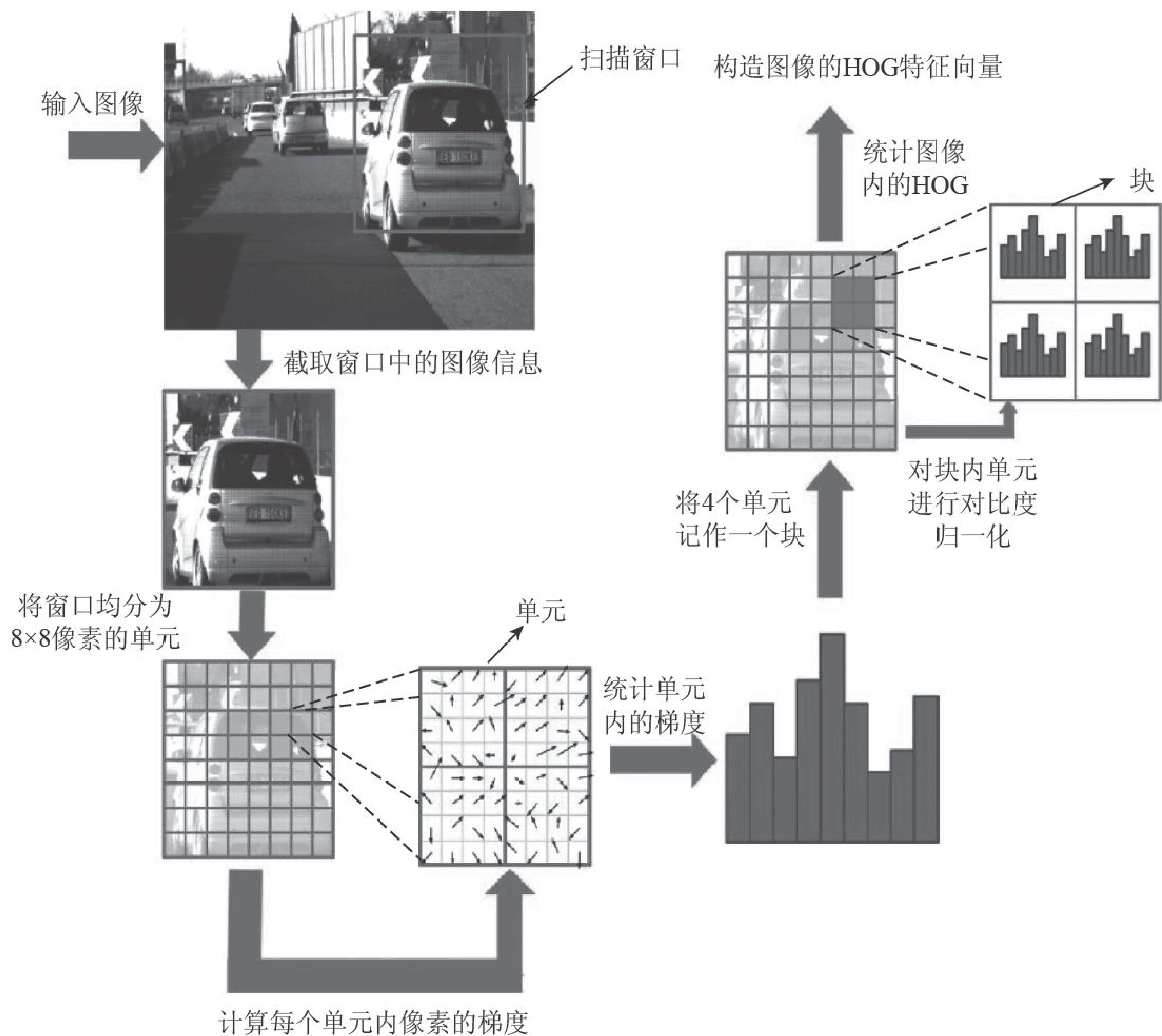


图6-14　提取和构造图像HOG特征的主要步骤

6.4　实验结果与分析

本书进行了两幅样本图像的HOG特征描述。实验一以图6-15作为样本图像，手工标定目标区域，用矩形框表示。整幅图像的HOG特征图，如图6-16所示。标定图像块的HOG特征图，如图6-17所示。

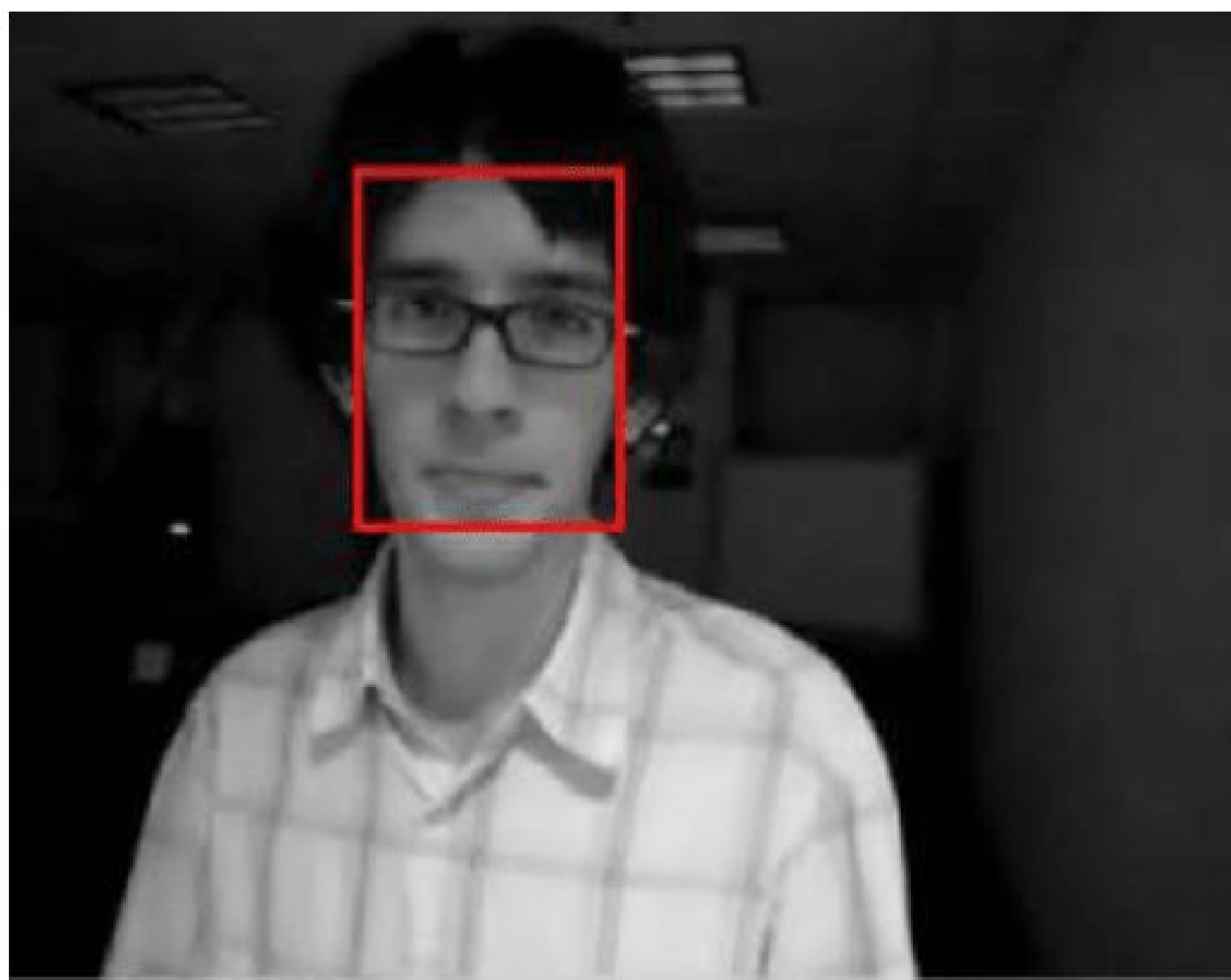


图6-15 样本图像

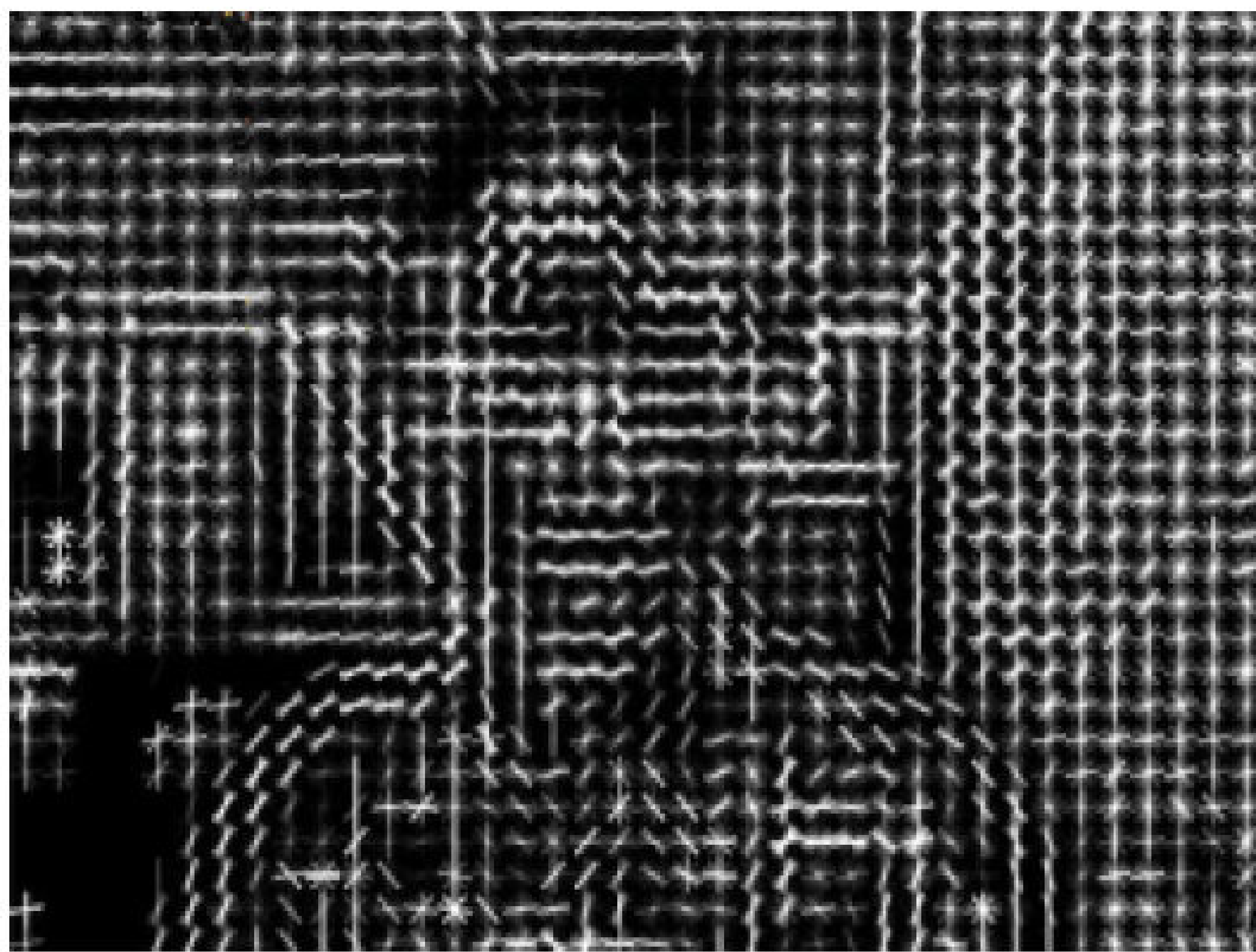


图6-16 图像HOG特征图

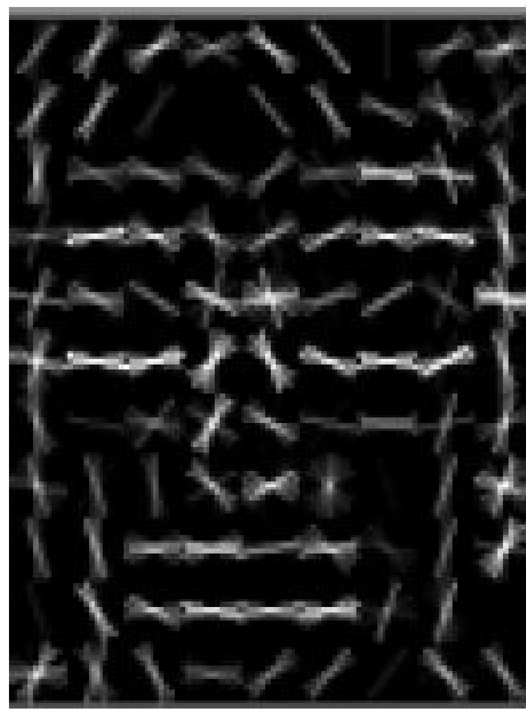


图6-17 目标图像块的HOG特征图

实验二以图6-18作为样本图像，选取该图像中的一个目标图像块，如图6-19所示，手工标定目标区域，用矩形框表示。该图像块的HOG特征图，如图6-20所示。

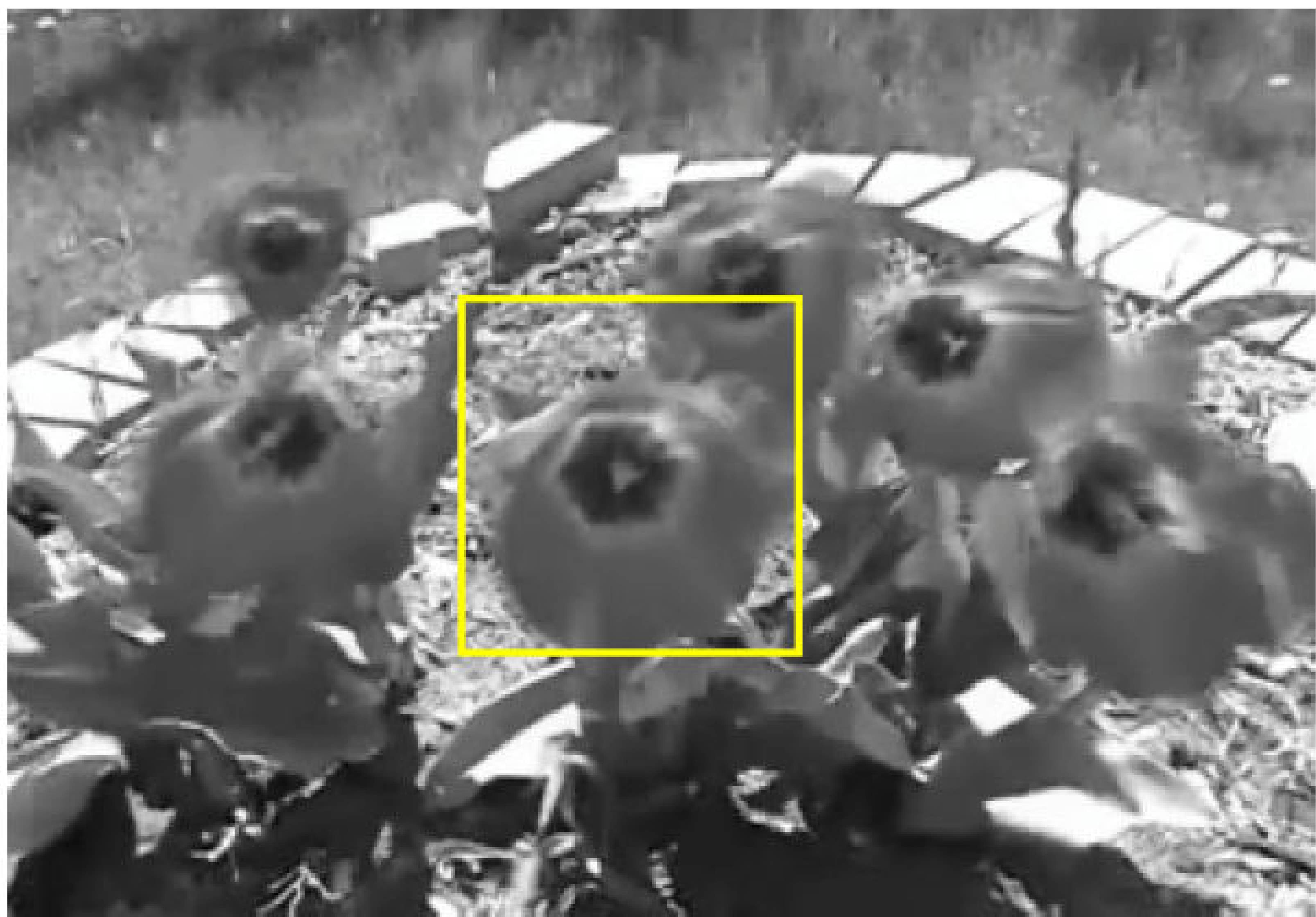


图6-18 样本图像

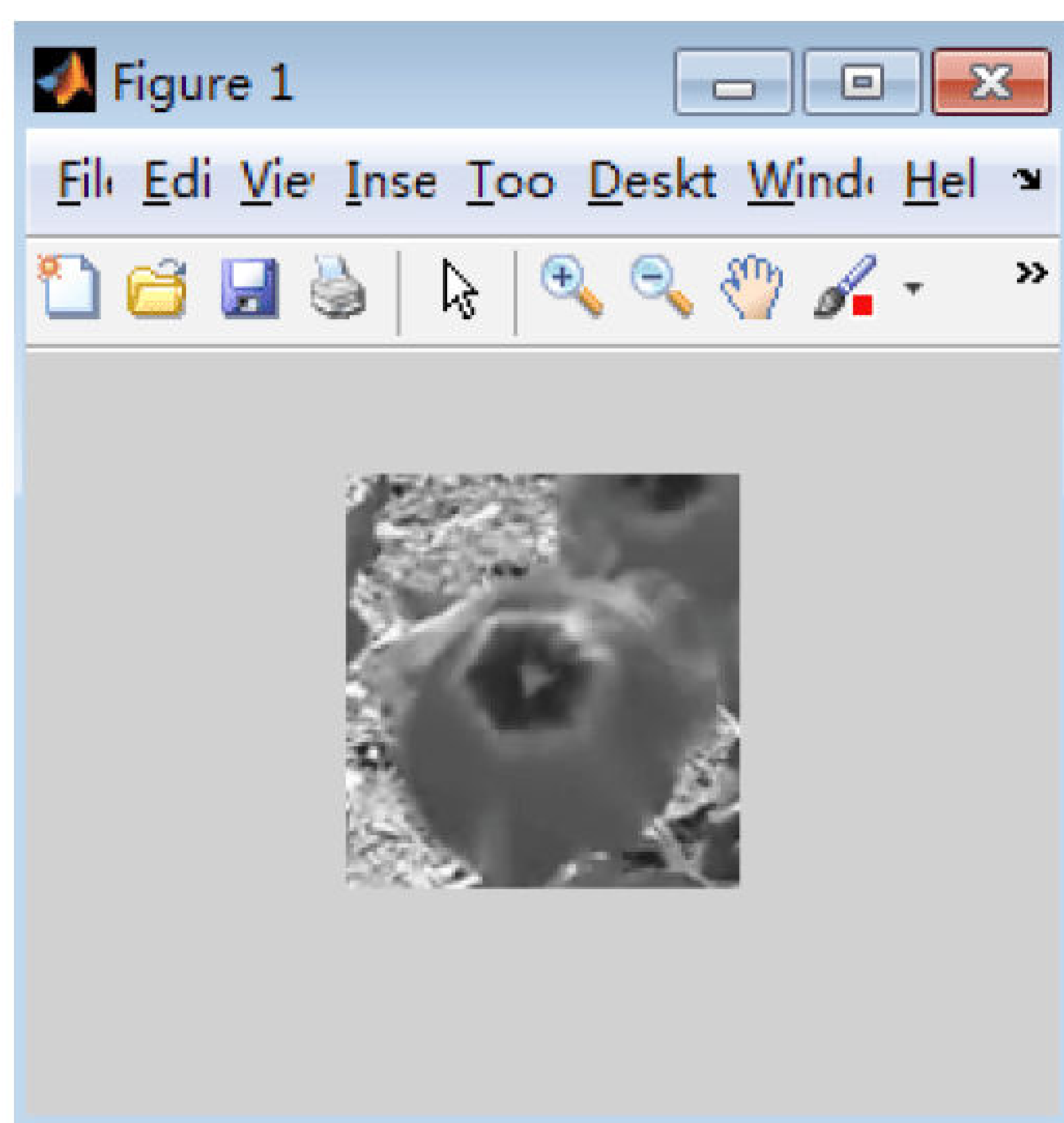


图6-19　选取图像块

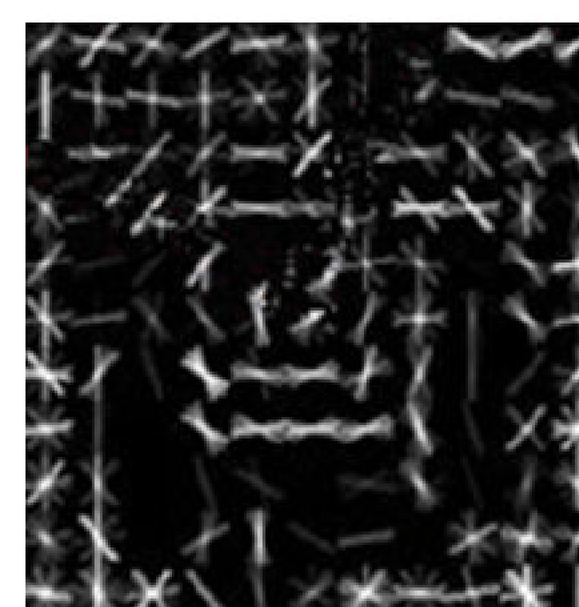


图6-20　图像块的HOG特征图

第7章

SVM机器学习算法



7.1 支持向量机(SVM)的原理

支持向量机(Support Vector Machine，SVM)是Vapnik[200]于1995年在统计学习理论研究和结构风险最小的前提下提出的一种机器学习分类算法。该算法的最大优点在于能有效解决小样本、非线性及高维模式等的识别分类问题，如今已成为通用的且性能良好的分类算法。

小样本：在实际应用中，对任何算法而言，选取的样本越多，最终的实验效果就会越好，但SVM算法达到同样的效果所需要的样本数会相对较少。

非线性：该点为SVM的精髓，主要是指SVM通过惩罚因子和核函数的参数设置，有效处理非线性可分的样本数据。

高维模式：SVM分类器非常简洁，仅用样本数据中“支持向量”的少量样本，即便样本维数高，存储和计算也不会特别复杂。

SVM算法的基本思想是：对于线性可分情况，寻找一个最优分类超平面，其在保证两类样本正确线性区分的前提下，使分类间隔最大。对于非线性情况，SVM通过一个核函数对原始数据升维运算，映射到另一个新的高维特征空间上，使得样本在该空间上线性可分，由此建立的最优分类超平面在原始空间内是一个超曲面。

7.2 SVM分类器的数学模型

7.2.1 线性可分情况下的SVM模型

SVM分类方法适用于线性可分情况下的分类，如图7-1所示。

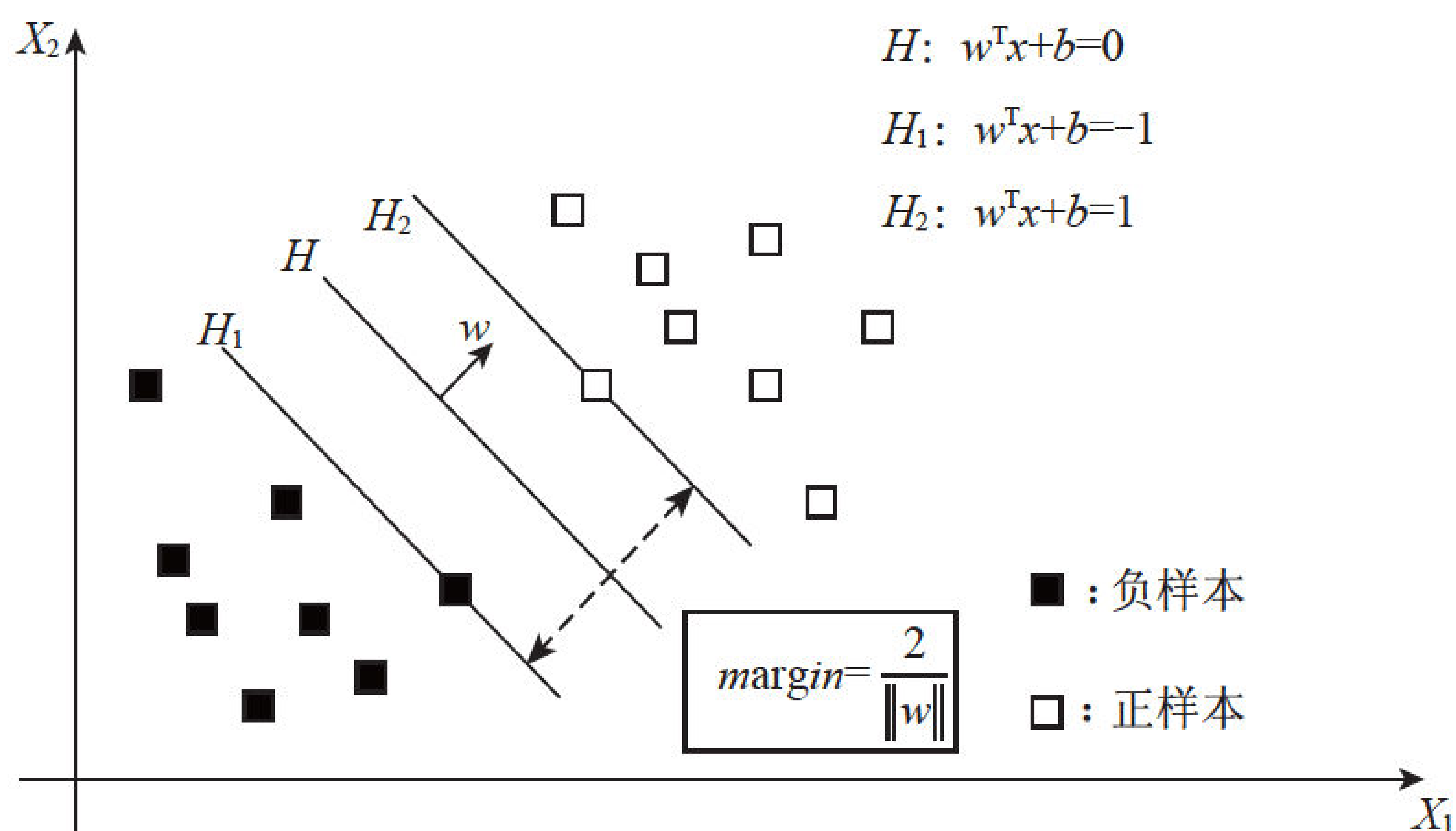


图7-1　二维线性可分问题示意图

两类训练样本由空心方块和实心方块表示。H为分类线，即样本集合中所有的两类数据均可以被H正确划分。直线H_1、H_2都与H平行，且两类数据中离分类线最近的样本均在直线H_1、H_2上。H_1、H_2之间的距离就是分类间隔(Margin)，间距为$2/\|w\|$。如果H既能满足两类数据无错误地分开，且能使分类间隔最大，它就是寻找的最优分类线，位于H_1、H_2上的样本即为支持向量。

对于一个训练数据集合(x_i，y_i)，i=1，2，…，n，$x_i \in R^d$，y_i∈{□1，1}即x_i为d维特征向量，y_i∈{□1，1}表示类别，最优超平面H的方程表示为$w^T x+b=0$，则d维空间的线性判别函数为$g(x)=w^T x+b$。w是SVM超平面法向量，也就是分类方程的系数。不失一般性，两类数据满足可分的情况有以下关系

$$\begin{cases} w^T x_i + b \geqslant 1，当 y_i = 1 时 \\ w^T x_i + b \leqslant -1，当 y_i = -1 时 \end{cases} \tag{7-1}$$

将式(7-1)归一化，即两类样本满足

$$y_i\left(w^T x_i + b\right) \geqslant 1 \tag{7-2}$$

则我们可以得到H_1与H_2之间的距离，即

$$\rho = \left|d_{H_1} - d_{H_2}\right| = \left|\frac{w \cdot 0 + b - 1}{\|w\|} - \frac{w \cdot 0 + b + 1}{\|w\|}\right| = \frac{2}{\|w\|} \tag{7-3}$$

那么，最后超平面的求解就是要求式(7-3)的最大化等价于最小化2/||w||，于是由寻找最优超平面转化为一个约束优化问题，即

$$\begin{aligned} & \min_{w,\ b} \quad \varphi(w) = \frac{1}{2}\|w\|^2 \\ & s.t. \quad y_i\left(w^T x_i + b\right) \geqslant 1，i=1，2，\ldots，N \end{aligned} \tag{7-4}$$

该问题实际是二次规划(Quadratic Programming，QP)问题，又可以行域为凸集，故这是一个凸二次规划问题，因此可以转化为拉格朗日(Lagrange)问题处理，求解最优参数w、b，同时也就得到了最优分界面，然后可以对未知样本进行分类。

构造Largange函数，即

$$\begin{aligned} & \min_{w,\ b,\ \alpha} L(w,\ b,\ \alpha) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{n} \alpha_i \left[y_i \left(w^T x_i + b \right) - 1 \right] \\ & \alpha_i \geqslant 0,\ i=1,\ 2,\ \ldots,\ n \end{aligned} \tag{7-5}$$

其中α是Largange乘子。

(1) 首先不考虑α，分别对w、b求偏导，即

$$\begin{aligned} & \frac{\partial L}{\partial w} = 0 \Rightarrow w = \sum_{i=0}^{n} a_i y_i x_i \\ & \frac{\partial L}{\partial b} = 0 \Rightarrow \sum_{i=0}^{n} a_i y_i = 0 \end{aligned} \tag{7-6}$$

把式(7-6)的计算结果带入L，得到原问题的对偶(Dual)问题，即

$$\begin{aligned} & \max_{\alpha} \ L_\alpha(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i,\ j=1}^{n} \alpha_i \alpha_j y_i y_j x_i^T x_j \\ & s.t. \quad \sum_{i,\ j=1}^{m} \alpha_i y_i = 0，\alpha_i \geqslant 0，i=1，2，\ldots，m \end{aligned} \tag{7-7}$$

它的具体推导过程为

$$
\begin{aligned}
L(w, b, \alpha) &= \frac{1}{2}\|w\|^2 - \sum_{i=1}^{n}\alpha_i\left[y_i\left(w^T x_i + b\right) - 1\right] \\
&= \frac{1}{2}w^T w - \sum_{i=1}^{n}\alpha_i y_i w^T x_i - \sum_{i=1}^{n}\alpha_i y_i b + \sum_{i=1}^{n}\alpha_i \\
&= \frac{1}{2}w^T \sum_{i=1}^{n}\alpha_i y_i x_i - \sum_{i=1}^{n}\alpha_i y_i w^T x_i - \sum_{i=1}^{n}\alpha_i y_i b + \sum_{i=1}^{n}\alpha_i \\
&= \frac{1}{2}w^T \sum_{i=1}^{n}\alpha_i y_i x_i - w^T \sum_{i=1}^{n}\alpha_i y_i x_i - \sum_{i=1}^{n}\alpha_i y_i b + \sum_{i=1}^{n}\alpha_i \\
&= -\frac{1}{2}w^T \sum_{i=1}^{n}\alpha_i y_i x_i - \sum_{i=1}^{n}\alpha_i y_i b + \sum_{i=1}^{n}\alpha_i \\
&= -\frac{1}{2}w^T \sum_{i=1}^{n}\alpha_i y_i x_i - b\sum_{i=1}^{n}\alpha_i y_i + \sum_{i=1}^{n}\alpha_i \\
&= -\frac{1}{2}\left(\sum_{i=1}^{n}\alpha_i y_i x_i\right)^T \sum_{i=1}^{n}\alpha_i y_i x_i - b\sum_{i=1}^{n}\alpha_i y_i + \sum_{i=1}^{n}\alpha_i \\
&= -\frac{1}{2}\sum_{i=1}^{n}\alpha_i y_i (x_i)^T \sum_{i=1}^{n}\alpha_i y_i x_i - b\sum_{i=1}^{n}\alpha_i y_i + \sum_{i=1}^{n}\alpha_i \\
&= -\frac{1}{2}\sum_{i=1, j=1}^{n}\alpha_i y_i (x_i)^T \alpha_j y_j x_j - b\sum_{i=1}^{n}\alpha_i y_i + \sum_{i=1}^{n}\alpha_i \\
&= \sum_{i=1}^{n}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{n}\alpha_i\alpha_j y_i y_j x_i^T x_j
\end{aligned}
\tag{7-8}
$$

(2) 求对偶问题的最优化，即求α的极大值。从公式(7-8)得到

$$
\begin{aligned}
&\max \sum_{i=1}^{n}\alpha_i - \frac{1}{2}\sum_{i,\ j=1}^{n}\alpha_i\alpha_j y_i y_j x_i^T x_j \\
&s.t.\ \alpha_i \geqslant= 0，i=1，2，\dots，n \\
&\sum_{i=1}^{n}\alpha_i y_i = 0
\end{aligned}
\tag{7-9}
$$

可求得唯一解$\alpha^*_i=(\alpha^*_i, \cdots, \alpha^*_n)$，则$\alpha$所对应的训练样本$x_i$即为支持向量。由$w=\sum_{i=1}^{n}\alpha_i y_i x_i$，求$w^*$值；再由$b^* = -\dfrac{\max_{i:y_i=-1} w^{*T}x_i + \min_{i:y_i=1} w^{*T}x_i}{2}$，求$b^*$值。最终，得到$f(x)=\mathrm{sgn}\left[\left(\omega^* \cdot x\right)+b^*\right]=\mathrm{sgn}\left[\sum_{i=1}^{n}a_i^* y_i\left(x_i \cdot x\right)+b^*\right]$，该函数就是所求的分类决策函数。

7.2.2 线性不可分情况下的SVM模型

对于大多数的实际问题，完成两类样本数据的线性分开是非常困难的。我们尽量寻找一个分类平面使尽可能少的样本点错分，可看作该样本能近似线性可分。数学上，利用松弛变量$\xi_i \geqslant 0$，$i=1$，2，…，n，使平面$w^T x+b=0$满足

$$\begin{cases} w^T x_i + b \geqslant 1-\xi_i，当 y_i=1时 \\ w^T x_i + b \leqslant -1+\xi_i，当 y_i=-1时 \end{cases} \tag{7-10}$$

使得式(7-10)具有一定的容错性，则相应的目标函数变为

$$\begin{aligned} &\min_{w,b,\xi} \quad \phi(w,\ b,\ \xi)=\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\xi_i \\ &s.t. \quad y_i\left(w^T x_i + b\right) \geqslant 1-\xi_i \\ &\qquad i=1，2，\ldots，n \end{aligned} \tag{7-11}$$

其中C为惩罚因子，C越大，惩罚越大，表明我们越重视错分点，越不愿意放弃。在整个求解过程中，惩罚因子C设置为一定值，需要参数寻优。$\sum_{i=1}^{n}\xi_i$是惩罚函数，代表错划程度，表示所有错分的样本点到相应正确分界面的距离。只有“错分点”才有松弛变量与之相对应，其他正确可分样本点都为0。

上述情况将线性不可分问题“强行地”看作线性可分进行处理，引入松弛变量和惩罚因子的目的，就是对错误分类的样本进行“包容”。

同样地，引入乘子α、β，构建Lagrange函数，并求解这个具有线性约束的二次规划问题，即

$$\begin{aligned} &\max_{\alpha,\beta}\min_{w,b,\xi} L_p = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}\xi_i - \sum_{i=1}^{N}\alpha_i\left[y_i\left(w^T x_i + b\right)-1+\xi_i\right] - \sum_{i=1}^{N}\beta_i\xi_i \\ &s.t. \quad 0 \leqslant \alpha_i，0 \leqslant \beta_i \end{aligned} \tag{7-12}$$

把上述问题转换为对偶问题，即

$$\begin{aligned} &\max_{\alpha} L_p = \sum_{i=1}^{N}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{N}\alpha_i\left(y_i y_j x_i x_j\right)\alpha_j \\ &s.t.\ 0 \leqslant \alpha_i \leqslant C \\ &\qquad \sum_{i=1}^{N}\alpha_i y_i = 0 \end{aligned} \tag{7-13}$$

可以看到，这种处理与线性可分情况的唯一区别是对α设定了一个上限。而且，对于非线性情况，该方法同样适用，从而使SVM算法对于模式识别问题具有通用性。

7.2.3　非线性可分情况下的SVM模型

对于最优分类面是非线性函数的情况，可通过一个非线性变换为一个新特征空间的线性问题求解。但实现这种方法要解决两个问题：一是在概念上如何在高维空间寻找一个推广好的最优分类超平面；二是在技术上如何处理高维空间中复杂的计算问题。如果数学上可以找到一个函数，K：$(R^N, R^N) \to R$ 使得$K(x_i, x_j)$等于x_i、x_j在高维特征空间中的映射之内积，即$K(x_i, x_j)=\Phi(x_i) \cdot \Phi(x_j)$，其中$\Phi(x)$为$x$映射到高维线性空间$x'$，则$K(x_i, x_j)$称为核函数。那么用核函数$K(x_i, x_j)$直接代替最优分类平面中的点积$x_i$、$x_j$，在计算的复杂程度几乎不变的情况下，就可以把原特征空间升维到另一高维特征空间。训练完成后的决策函数变为$f(x)=\mathrm{sgn}\left[\sum_{i=1}^{l}\alpha_i y_i k(x, x_i)+b\right]$。

可以看到，支持向量机分类与传统分类方法的处理思路截然相反，该算法不是按常理将原特征空间降维，而是设法升维，通过高维空间把平面上本身不好分的原数据分开，完成非线性不可分情况到线性可分或接近线性可分情况的转换。由于通过核函数进行升维，向量内积运算方式被改变，并没有提高问题的计算复杂度。在图7-2中，两类数据样本在二维空间中线性不可分，可升维到三维空间完成原数据的分类。

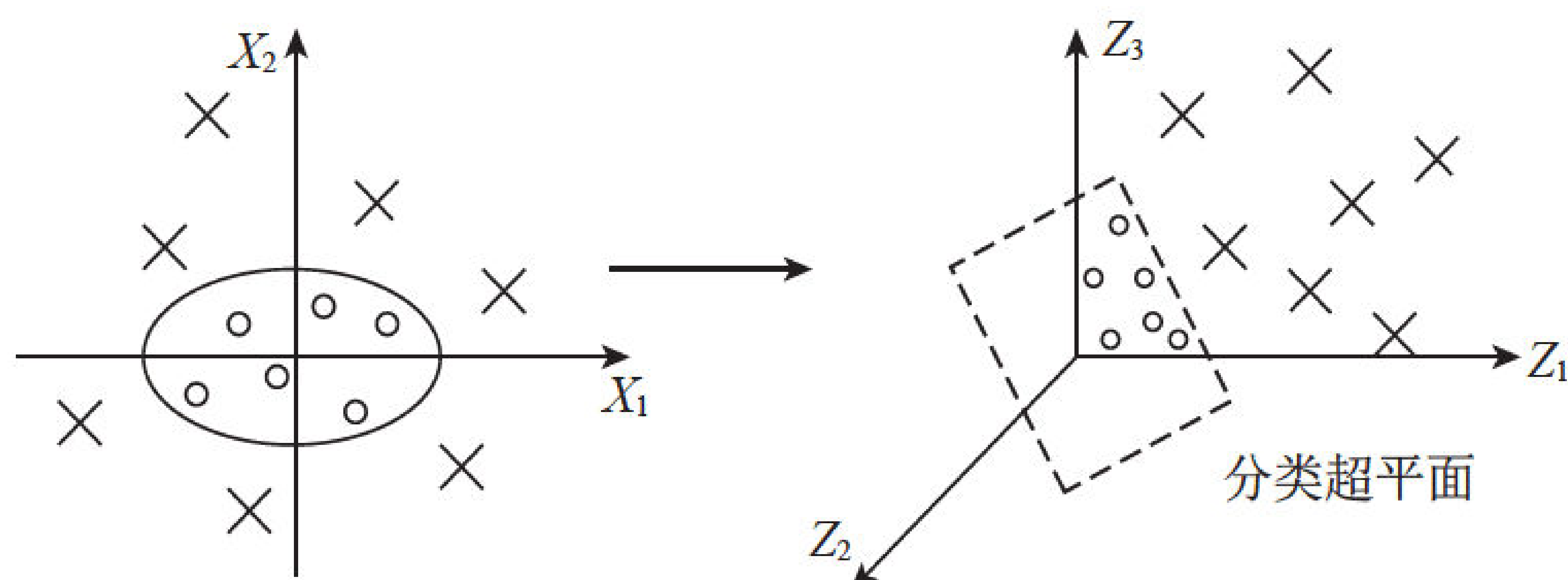


图7-2　非线性可分通过升维达到线性可分的示意图

常见的核函数有

(1) 线性核函数：$K(x_i, x_j)=x \cdot x_i$；

(2) 多项式核函数：$K(x, x_i)=[(x \cdot x_i)+1]^q$，其中$q$的范围通常取1～10；

(3) 径向核函数：$K(x, x_i)=\exp\left(-\frac{|x-x_i|^2}{\sigma^2}\right)$，其中$\sigma$的范围通常取0.001～0.006；

(4) S型核函数：$K(x, x_i)=\tanh[a(x \cdot x_i)+\mathrm{b}]$。

本书采用的是线性核函数，线性核函数的运算速度比较快，会提高系统的实时性，使平均运行速度提高。因为使用线性核函数可以直接将待检测图像与SVM训练模板进行卷积操作，而不需要将模板特征和待检测图像特征映射到核函数对应的特征空间中。通过迭代方式可以将SVM训练的过程理解为对含约束条件的二次优化问题求解的过程[201]，此处设一定的容许误差以及最大的迭代次数，以便在适当条件下使该算法的计算过程停止。

7.3　SVM分类器设计

SVM分类器通过寻找一个最优分割超平面，其在保证两类样本正确区分的前提下，使分类间隔最大。选择样本集$\{x_i, y_i\}_{i=1}^{l}$，其中，$y_i \in \{$□$1, 1\}$为各样本标签。最优超平面函数的确定可通过含约束条件的二次规划问题(Quadratic Programming)求解。分割超平面的表达公式为

$$f(x)=\sum_{i=1}^{l}\alpha_i y_i k(x, x_i)+b \tag{7-14}$$

其中，$\{x_i, y_i\}_{i=1}^{l}$为一组训练样本，x为一个新的测试用例。$k(x, x_i)$为核函数，用来将SVM的分类超平面从线性不可分情况转化为非线性可分情况。判别函数$f(x)$用来区分x的类别。若α_i对应的训练样本x_i为支持向量，构建超平面的关键就是求解非零数组α_i。

在进行目标跟踪时，向分类器输入一个候选目标x^T，则对应的判别输出为

$$f(x_T)=\sum_{i=1}^{l}\alpha_i y_i k(x_T, x_i)+b \tag{7-15}$$

最后，使分类器具有最大输出值的x^T确定为跟踪目标，即$\left\{x_T \,\middle|\, \max_{x_T\in\Theta}\sum_{i=1}^{l}\alpha_i y_i k(x_T, x)\right\}$，其中$\Theta$表示候选目标空间。

7.4　HOG和SVM的目标外观模型

7.4.1　目标外观模型的建立

Dalal-Triggs所设计的行人检测器通过单一的HOG特征滤波器来表示目标类别，并利用检测窗口依次滑动搜索图像所有的位置来检测目标，它在2006年的PASCAL VOC目标检测比赛上获得冠军。本书就是通过Dalal-Triggs检测器提取目标HOG特征，用支持向量机(SVM)分类器训练，得到目标HOG模板。

该检测器相当于一个分类器，将图片、图片的像素坐标以及图片尺度作为输入。这个分类器通过滤波运算判定在给定尺度的窗口内是否存在目标实例。模型是一个单一的滤波器，可以通过下式计算得分

$$\text{score}=\beta \cdot \phi(x) \tag{7-16}$$

式中β为滤波器，是模型参数的向量形式；x是输入图像，指定了像素点坐标以及尺度大小，$\phi(x)$表示特征向量。

本章在初始帧中随机选择大量含目标区域以及含背景区域的等大图像块，分别提取HOG特征，标记为正负样本，并将这些带有标签的训练样本输入SVM分类器，使分类器具有一定的分类能力，输出相应的HOG特征模板(即分类器权值)，再经过与SVM训练模板卷积操作，即可获得目标准确的参数化外观模型。具体实现细节如下：

(1) 根据第一帧已标注的目标位置进行正负样本的采样，提取两类正负样本的HOG特征，从而建立特征空间下的初始化训练集合。对于HOG特征，根据本书6.2节对于HOG特征提取的原理及算法的描述，本书采用8×8个像素大小的小细胞单元Cell，由4个Cell组成的一个局部区域块Block，每个Cell内有9个梯度方向的维度，每个Block以8个像素宽度为步长对输入图像进行扫描后可以提取出每个Block内的HOG特征，每个Block内有4个9维的梯度特征向量，这样就可以提取出每个样本中高维的HOG特征向量。

(2) 对所有样本给予相应的样本标签。例如，标记正样本为1，负样本为−1。SVM通过对样本的学习找到特征向量样本中的支持向量，建立最优分类超平面，得到具有一定分类能力的分类器，再将当前视频中运动目标的检测结果作为测试样本，并提取它的特征向量，输入SVM即可训练得到目标的参数化外观模型。

7.4.2 SVM分类器的训练

SVM分类器的训练采用一个简单易用、快速有效的SVM软件包——LibSVM函数库。该函数库提供了源代码，方便修改并有可执行文件，可在其他操作系统上应用，可以让非专业用户更加方便快捷地使用SVM这个统计学习工具，而无须考虑其内部复杂的数学模型和运算过程。它对SVM所涉及的参数调节较少，LibSVM库中有各种SVM表达公式，同时包含4种常用的核函数及其默认参数，可适用于很多分类情形，并提供了交互检验(Cross Validation)的功能来分析分类识别的正确率。

下面是利用LibSVM来训练SVM分类器的过程。本书讨论的学习样例均为图像块(Patch)，即每个图像块对应一个机器学习的实例。采用图像块描述目标对象，其结果直观、灵活，且抗变换能力强。图像块本身能达到抗平移变换的效果，在图像块上做适当的特征提取(如颜色或者梯度直方图)能够抵抗一定的仿射变换与变形变换；而对于旋转变换，可以通过多个不同朝向的图像块来描述。由于对图像的基本操作(如特征提取)都在图像块上进行，因此基于图像块的机器学习过程比基于复杂物体部分的方法更加简单快速，特别适用于在线学习问题，满足了目标对象跟踪的实时性要求。

以红花视频中的一个图像块为例，首先根据第一帧已标注的该目标位置进行正负样本的随机采样，各取得表示目标的50个正样本以及表示背景的100个负样本，并给予相应的样本标签。将这些训练样本数据转换为LibSVM要求的文件格式后，利用svm-train.exe程序训练分类器，如图7-3所示。

alpha <1x1 struct>

Field	Value	Min	Max
Parameters	[0;4;3;0.0066;0]	0	4
nr_class	2	2	2
totalSV	17	17	17
rho	0.9997	0.9997	0.9997
Label	[0;1]	0	1
ProbA	-4.2401	-4.24...	-4.24...
ProbB	0.3343	0.3343	0.3343
nSV	[1;16]	1	16
sv_coef	<17x1 double>	-0.12...	0.2088
SVs	<17x1 double>		

图7-3　训练输出结果

svm-train.exe程序的输出结果：totalSV为支持向量的总数目；Label设置分类中的类别标签；rho为判别函数的偏置项b；nSV是每类支持向量的个数；sv_coef为支持向量对应的权重系数wi；SVs是相应的支持向量。

下面为分类器训练和验证的过程。

```
K>> a = [pos{ii} : neg {ii}] :
K>> a_label=a (:, end) :
K>> a_inst=a (:, 1: end-1):
K>> model = svmtrain (a_label, a_inst) :
K>> [predict_label, accuracy] = svmpredict (a_label, a_inst, model):
Accuracy = 100% (150/150) (classification)
```

由此可以看出，通过训练样本训练出的分类器的分类正确率达到100%，验证了利用SVM分类对运动目标进行识别的可行性。对测试样本数据的验证之所以能达到100%，是因为测试样本的数据与训练样本的数据相差不大。实际中，对通过运动检测出的目标进行识别时，因考虑到各种因素的影响将不会达到100%的正确率。SVM分类器训练好后，保存在Model的文件中，以后再利用该分类器进行目标识别时，只需加载该文件，即可获得分类器模型。

第8章
可变形模板的多目标跟踪模型



由于在不同的环境中单个个体的形态差异很大，而且不仅有光照和视角的变化，还有目标的非刚性变形，各种复杂因素使得运动目标的视觉外观变化诸多，现有的模型不能及时适应物体的运动，特别是非刚性物体的运动，因此要研究一个合适的建模方法来满足目标检测的需求。

8.1　可变形模板

可变形模板单独对各部件本身进行描述，并对不同部件之间的关系进行建模，进而通过组合来描述整体模板，从而比单一模板灵活，能处理不同的局部遮挡和结构变化。可变形模板及模型用空间相关的部件模型的集合来表示目标，对每个部件(可变形)目标的外观建模、部件之间的空间关系是由部件位置的一些联合分布建模而成的。这个部件位置的常见的联合分布选取包括：

(1) 主元分析法(PCA)模型：使用一个低级高斯模型。具有代表性的是主动外观模型(Active Appearance Model，AAM)、主动形状模型(Active Shape Model，ASM)、局部约束模型等。[202]基于主元分析法(Principal Component Analysis，PCA)的模型被广泛应用在人脸或医学图像分析中。

(2) 图结构模型：使用部件间弹簧的集合，有代表性地使用树结构。基于图结构的模型通常应用在目标检测、姿态估计和手势识别的任务中。一般的训练数据不足以用来训练基于PCA的模型，这就是我们重视图结构的原因。

8.2　图结构模型

图结构本身就是一种可变形模板。基于部件的图结构模型其实是一个弹簧模型，模型的每一部分都通过弹簧连接。图8-1是人脸的弹簧模型。部分之间的关系将这个物体抽象为由弹簧相连的物体，各组成部分之间的相互运动也是在弹簧的相互作用下进行的。

本书对多目标间的空间位置关系给予重视，通过图结构的思想来构造相互关系模型，从而增加一定程度的空间结构约束，将可变形模板检测原理应用于多目标跟踪场景。

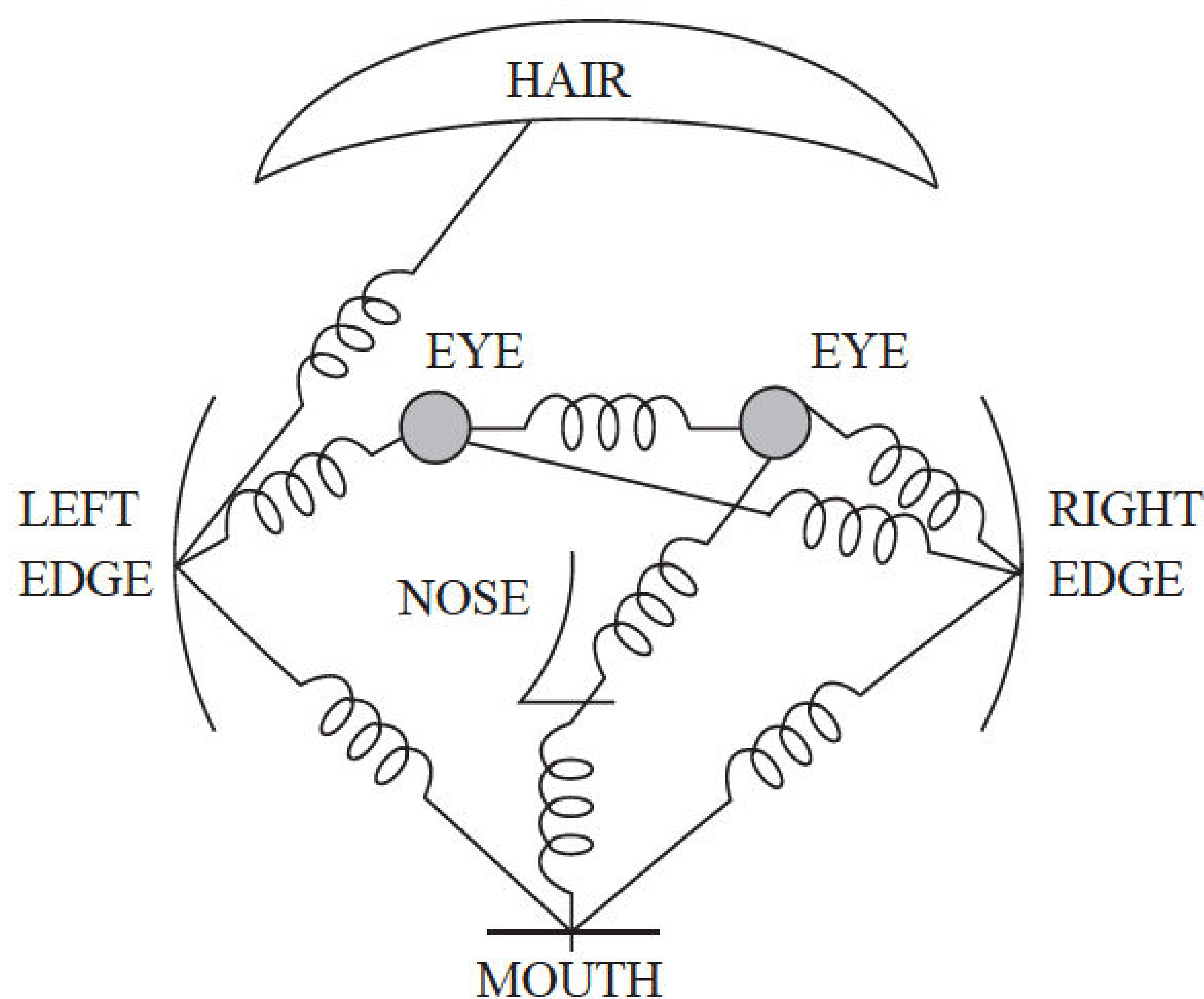


图8-1　人脸的弹簧模型

形式化表示中，基于部件(Part-Based)的可变形模型可以用无向图G=(V，E)来表示，V={v_1，v_2，…，v_3}代表n个部分，边(v_i，v_j)∈E代表两部分间的连接。物体在图像中的位置由其部分结构给出，即L={l_1，l_2，…，l_n}，l_i是第v_i个部分的位置。给定一幅图像，v_i被放置在图片中的l_i位置时，与模板的匹配程度为$m_i(l_i)$；v_i、v_j被分别放置在图片中的l_i、l_j位置时，模型的变化程度为$d_{ij}(l_i，l_j)$。因此，一幅图像相对于模型的最优配置，就是使其既能与每一部分的匹配度好，又能使部分间连接的变化程度小，即部分间的相对关系与模型尽可能相符。定义一个能量方程用来描述这两部分，与模型匹配最好的对象就是能使这个能量函数最大的数据对象。可以通过公式(8-1)描述最优配置(Configuration)。

$$L^* = \arg\max_L \left[\sum_{i=1}^{n} m_i\left(l_i\right) - \sum_{\left(v_i, v_j\right)\in E} d_{ij}\left(l_i，l_j\right) \right] \tag{8-1}$$

在之前的许多研究中，包含目标部件间最重要的空间关系由图结构手动设计。例如，用8部分来表示人体上半身的分层结构(上/下手臂、头部、躯干和手)。[203][204]为了使模型更加灵活，进一步将这些上半身的部分划分为更小的子部分。[205]然而，当设计的模型用于检测或跟踪任意目标时，这种用手动方法设计图结构模型的方式是不可行的。因此，任意目标的检测器通常使用通用的结构模型。由于物体组成部分存在空间上的位置关系，我

们需要考虑物体组成部分间的连通性，根据连通性结构的不同，模型结构有两种：①Star Shape(星形)，每一部分都与某一个中心部分相关联，其余部分间相互独立；②Tree(树形)，模型中的子节点只与其父节点存在关联。

8.2.1 图的基本概念

由于图(Graph)可以保留图像的结构以及目标区域之间的相互关系，因此，它可以作为一种非常重要并且有效的表示结构特征信息的描述工具。图像到图的转化是用图G中的顶点$v_i \in V$表示图像中的像素，用图的边e_{ij}表示像素特征之间的连接关系，其差异性或相似性由定义在边上的权重w_{ij}度量，从而将图像映射到加权图(Weighted Graph)。

提到树结构的图模型，我们应首先定义图。[206]

一个图G是由非空点集$V=V(G)$和V中元素的无序对的一个集合$E=E(G)$所构成的二元组，记为$V=[V(G), E(G)]$，简记为$G=(V, E)$。其中$V(G)=\{v_1, v_2, \cdots, v_n\}$称为图$G$的节点，$E(G)=\{e_1, e_2, \cdots, e_n\}$是$G$中边的集合。

在每一帧目标图像中，图是由节点和边所构成的完整集合，即$G=(V, E)$，本书中每一个目标物体都定义为图中的一个节点，目标间的空间相互关系为图中的边。下面是图G中一些相关的基本概念。

(1) 邻接节点：由边相连的两个节点，记为$v_i, v_j \in E(G)$；

(2) 环：两端节点相同的边；

(3) 平行边：方向相同的若干边；

(4) 阶：图G的节点个数；

(5) 度：与图G的任一节点v相连的边的个数，记为dv；

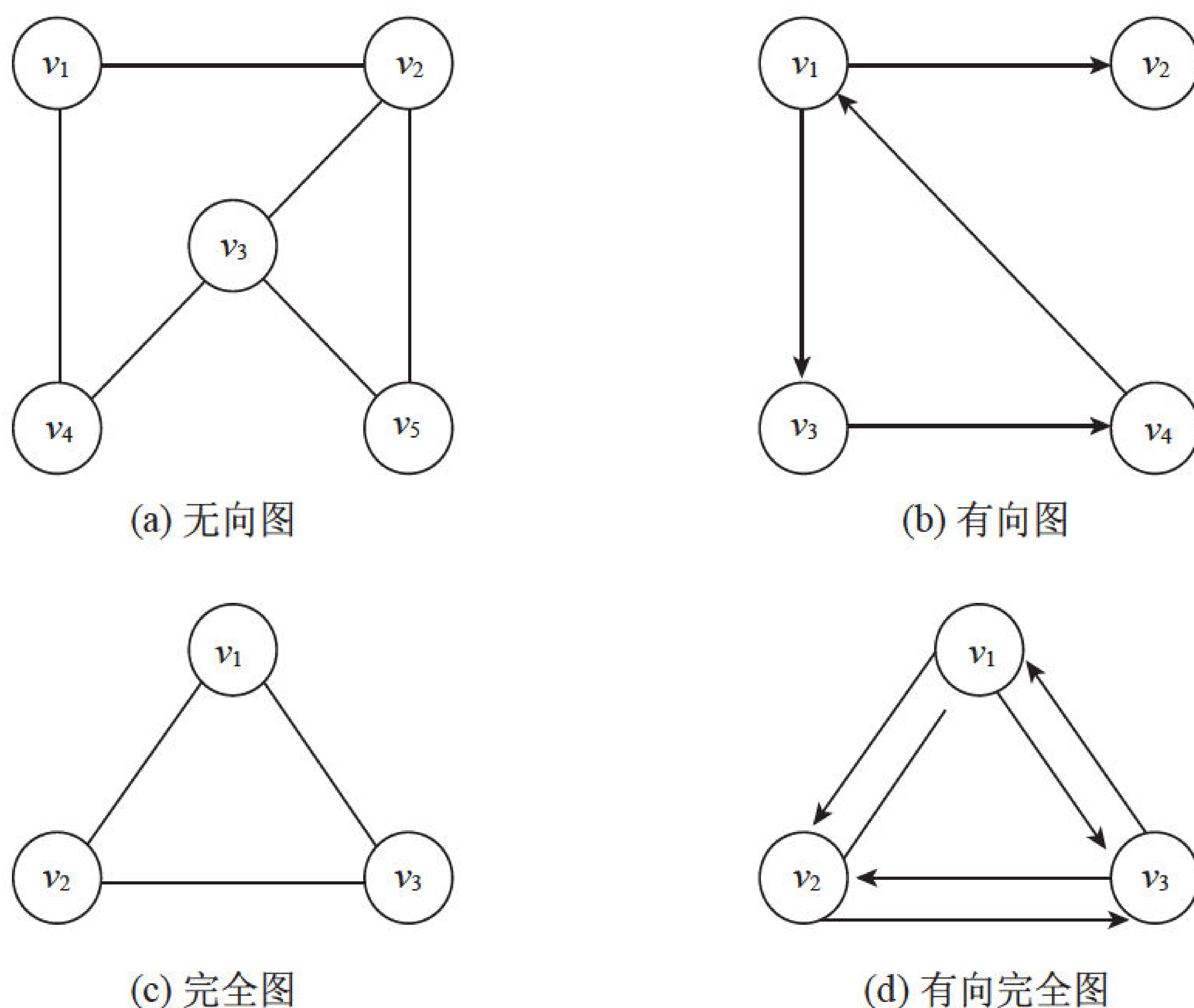
(6) 无向图：每条边都是无向边的图；

(7) 有向图：每条边都是有向边的图；

(8) 简单图：无环并且无平行边的图；

(9) 完全图：任意两节点都有边连接的简单无向图。

图8-2给出了无向图、有向图、完全图以及有向完全图的示例。



(a) 无向图　(b) 有向图

(c) 完全图　(d) 有向完全图

图8-2　各种图的示例

8.2.2　表示图像结构特征的图

由于选择图中不同的节点和不同的边，就可以组成不同的图结构，因此，图可以很好地对图像中目标物体的结构模型进行表征。图像处理任务中常用的图结构[209]有以下两种。

1. 赋权完全图

设图像含n个特征值，用$X=\{x_1, x_2, \cdots, x_n\}$表示，以点集中$x_i$作为图中节点$v_i$来构造完全图$G_x=(V_x, E_x)$，即$V_x=\{v_1, v_2, \cdots, v_n\}$，图中相应节点之间的边的权值表示特征点之间的距离，即$E_x=\{e_1, e_2, \cdots, e_m\}$，比如用欧式距离来表征，则$e_i = \|x_i - x_j\|$。

2. 最小生成树(Minimum Spanning Trees，MST)

若图结构树模型中含所有的n个节点和n□1条边，则该树模型称为图G的生成树(Spanning Trees)。若定义若干个权值分别赋予原图G中的各边，则形成G的加权无向连通图；在保证不出现回路的前提下，依次连接图中权值最小的边，最后构成的这个权值和最小的生成树就是图的最小生成树(MST)。

构造过程中需要满足两个条件：首先，在保证边不构成回路的前提下，寻求权值最小

的恰当边；其次，保证选取的n□1条边能连接图中所有的n个节点。若图中各边给定的权值都不等同，那么它有唯一结构的最小生成树模型。

8.3　构造最小生成树

赋权完全图强调任意两个顶点之间都有权值表征，顶点之间两两相互连接，信息表达丰富。然而，太多的信息处理起来比较麻烦，会增加问题的复杂性。最小生成树(MST)是图论中另一个用于描述图像结构特征的数学模型，由于其表达能提取模型中重要的结构，可有效降低模型的复杂度，近年来受到广大研究者的重视和肯定。

图8-4和图8-5所示的生成树$G1$和$G2$，都表示图8-3中G的生成树。$G1$和$G2$中所有的5个节点都相互连通且没有形成回路，但只有$G2$是G的最小生成树。G中最小的4个权重2、3、6、7分别对应(1，2)、(3，5)、(3，4)、(4，5)4条边，因为(3，5)、(3，4)、(4，5)3条边形成回路，故要去除相对权重最大的边(4，5)，而边(2，3)符合最小生成树的条件。无论G有多少棵生成树，最小生成树只能是图8-5所示的$G2$这唯一的形式。

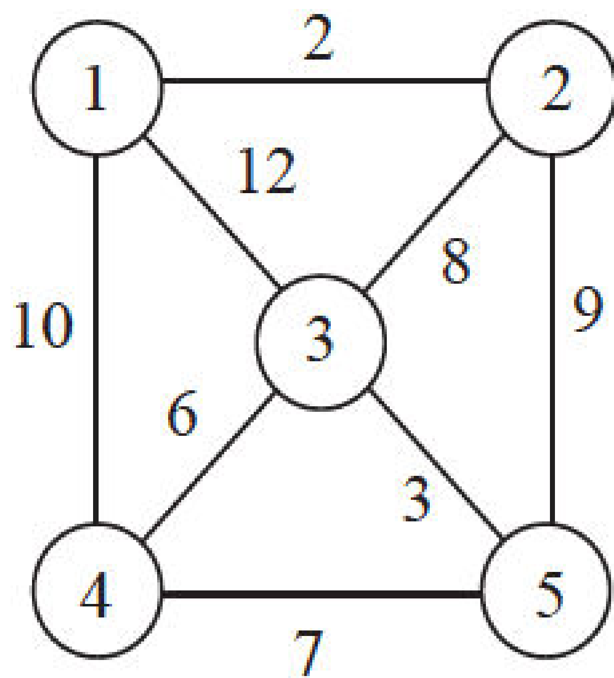


图8-3　G

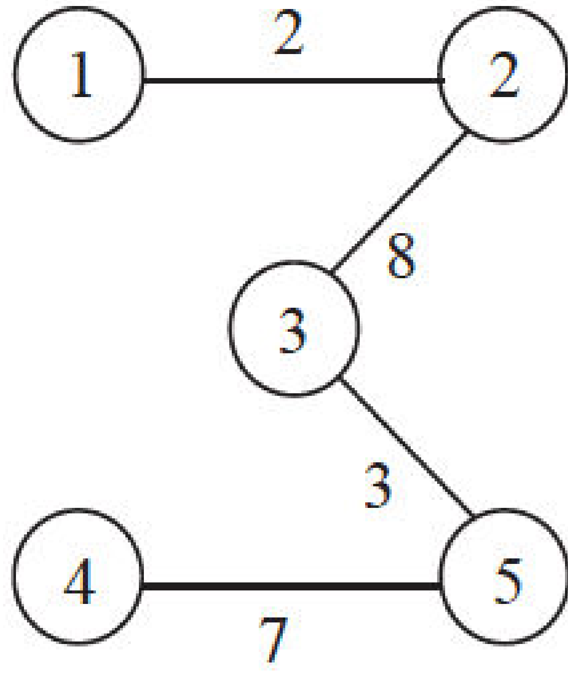


图8-4　G的生成树$G1$

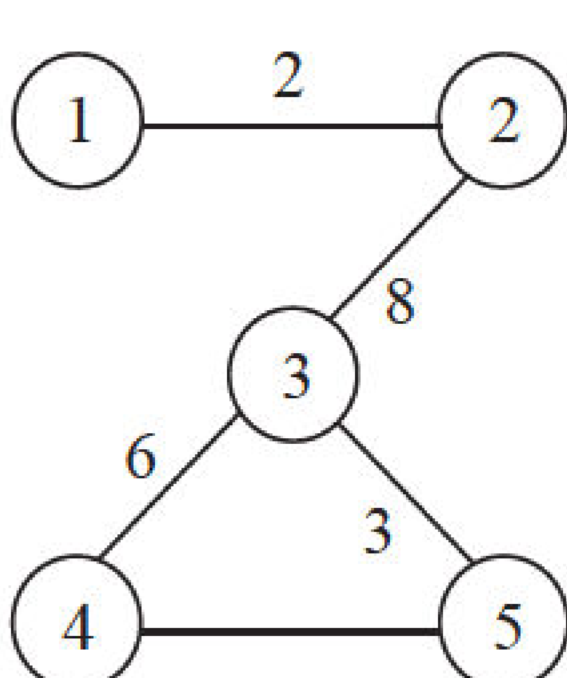


图8-5　G的生成树$G2$

最小生成树(Minimum Spanning Tree，MST)是数据结构中一个很重要的概念，经常被用于解决一些实际问题。常用的求解加权无向图最小生成树问题的方法有1957年提出的Kruskal算法和Prim算法。[210][211]各算法通过不同的选取策略，依次根据图中各边的权重构造出最小生成树，本书采用Kruskal算法。

下面介绍实现克鲁斯卡尔(Kruskal)算法的主要步骤。[212]

Step1：设图T=(V，E)初始化是含n个节点而边集为空的非连通图。

Step2：根据各边权值不同，把边集E中各边从小到大排序。

Step3：在E中选择权值最小的一条边，若该边的两个节点属于不同集合，则将其加入图中，且合并为一个集合；否则将构成回路，则放弃此边。

Step4：依次循环步骤3，直到T中包含所有的n个顶点和$n-1$条边，满足停止条件，完成最小生成树模型的构建。

图8-5中最小生成树$G2$的生成过程如图8-6所示，每次添加的边用粗线表示。

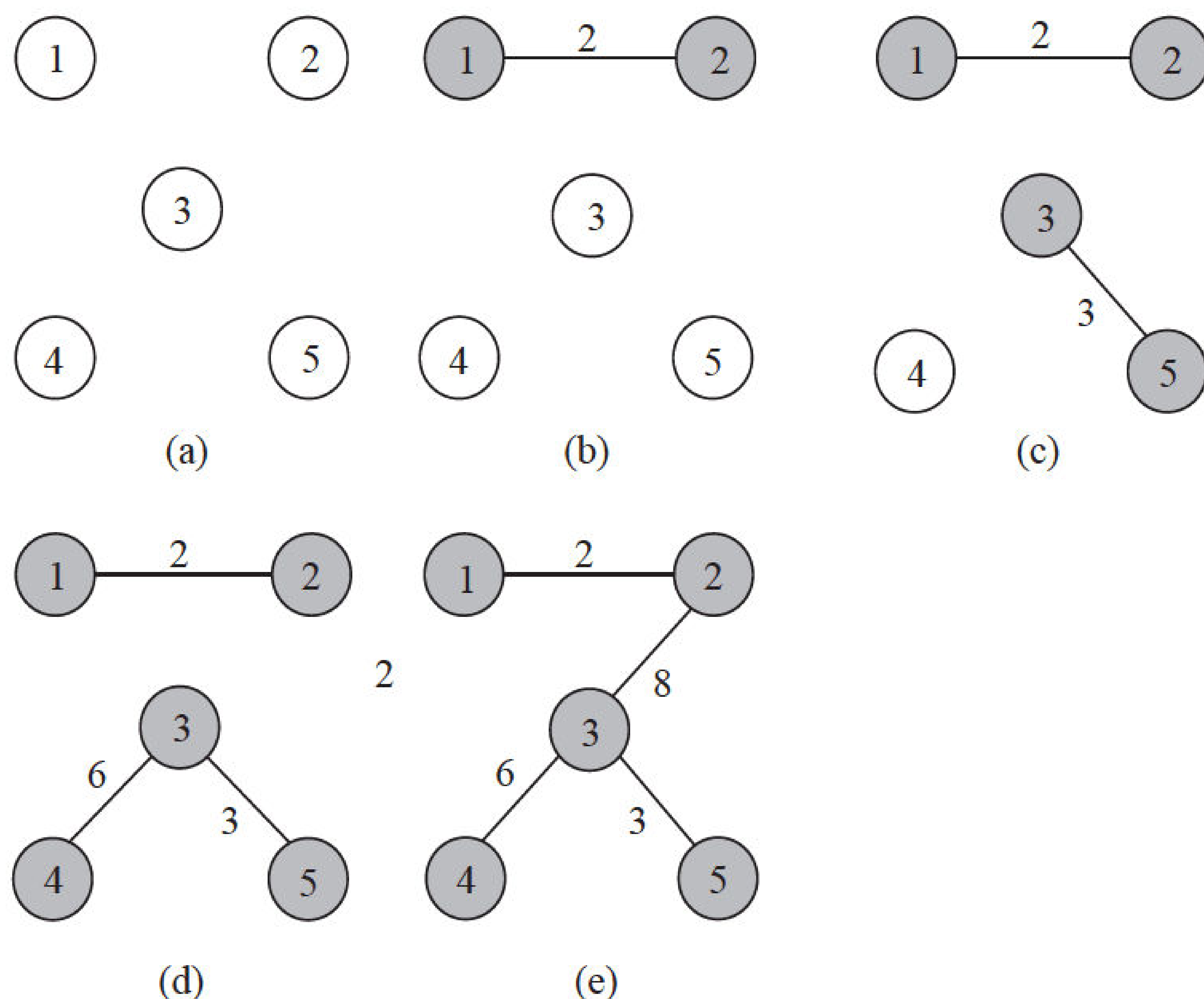


图8-6　最小生成树$G2$的生成过程

8.4　建立目标跟踪模型

可变形模板应用在模型各个部分的相对位移上能够提高模型的性能。首先，一个部分的遮挡可由另一个部分的位置来估计，并不会影响最后的检测结果，能够提高检测率；其次，有些部分所在的位置是基本相对固定的，而且各部分之间相互制约，可降低误判率。基于可变形模型的原理，提出最小生成树(MST)模型来建立各个部分之间的空间位置联

系，使整个检测过程的计算复杂度较小。

首先，我们手动标注初始帧中感兴趣的目标，描述表示目标$i \in V$(V代表我们正跟踪的目标的集合)的边界框为$B_i=\{x_i, w_i, h_i\}$，其中，中心位置为$x_i=\{x_i, y_i\}$，宽度为w_i，高度为h_i。从图像I中提取HOG特征，把目标边界框B_i内对应位置的特征连接起来，获得特征向量$\phi(I; B_i)$，即第i个目标的HOG特征描述。

然后，对所有目标$i \in V$定义一个无向图$G=(V, E)$，我们利用目标间的边$(i, j) \in E$来跟踪。在图模型中的边可看作弹簧，代表被跟踪目标间的空间约束。物体的某个实例的配置(Configuration)可以表示为$C=\{B_1, B_2, \ldots, B_{|V|}\}$，$B_i$表示第$i$个目标的位置(可以简单地将图片的配置理解为各部分的位置布局，实际上配置可以包含部件的其他属性)。

接下来，我们用多个被跟踪目标的空间配置C的得分来描述外观分数和结构变形花费。外观分数概括所有目标的图像特征和分类器权重之间的相似程度，通过计算窗口内的梯度方向直方图特征和一些权重的点积来获得；变形分数衡量部件间连接的变化程度，即部件偏离其理想位置的程度；也可以想象为弹簧的形变量，即配置对跟踪目标间的弹簧压缩或拉伸了多少。

数学上，基于可变形模板的多目标跟踪模型可用式(8-2)来描述。

$$s(C, \Theta)=\sum_{i \in V} \omega_i^T \phi(I; B_i)-\lambda \sum_{(i,j) \in E}\left\|(x_i-x_j)-e_{ij}\right\|^2 \tag{8-2}$$

式(8-2)就是配置C的得分。故检测窗口的得分可看作目标部分的外观匹配得分减去模型的结构变化花费。其中，$\phi(I; B_i)$表示图片I中B_i位置的HOG特征。参数$\grave{u}_i$是目标i在HOG特征中的线性权重，参数$\grave{u}_i$的初始化是第i个目标经SVM训练后获得的。这时SVM分类器基于第一帧中随机提取的正负样本训练而成。参数e_{ij}是第i和j目标间的相互关系参数，也可以理解为目标i和j之间弹簧的长度和方向的向量；其初始化基于目标的真实标注：$e_{ij} \leftarrow x_i - x_j$，$x_i$、$x_j$是目标$i$和$j$在第一帧中的位置。所有参数的集合表示为$\Theta=\{\omega_1, \ldots, \omega_{|V|}, e_1, \ldots, e_{|E|}\}$。我们把参数$\lambda$视为一个超参数，用来权衡外观分数和变形花费。实验中，固定λ=0.001。最后，我们确定图G的结构，也就是决定哪些目标由边互相连接。在理想情况下，我们应该对多目标构建一个完全连通的图G模型，用来约束多目标间的空间关系，但这会给后续的搜索检测过程带来大量复杂的优化问题。

因此，本书提出利用最小生成树模型来构建多目标间的空间相互关系模型，在有效表

示空间结构约束的同时，使整个检测过程的计算复杂度较小。最小生成树模型基于第一帧目标的标注，搜索所有可能的完全连接的树模型的集合，使其最小化$\sum_{(i,j)\in E}\left\|x_i - x_j\right\|^2$，其中$x_i$和$x_j$是目标$i$和$j$在第一帧中的位置。

最后，我们在所有可能的配置C中找到使等式获得最大值的最佳的目标配置输出，即各个目标的准确预测位置。

以红花视频图像序列为例，描述整个目标模型的构建过程。首先，读取视频序列的初始帧，如图8-7所示。



图8-7　初始帧图像

然后，手动标注视频序列初始帧中感兴趣的目标区域，用矩形框划分成图像块，可以看到目标被拆分为4部分，如图8-8所示。

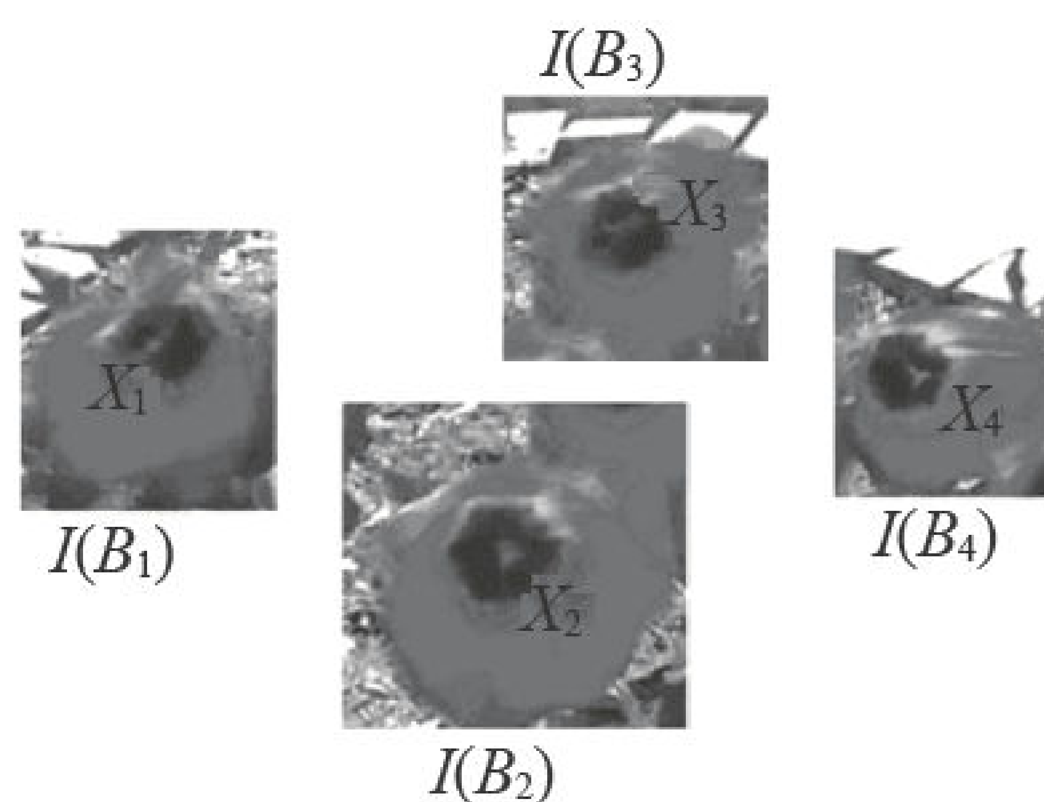


图8-8　目标图像块

图8-9是经过HOG特征提取后可以获得的4个特征块。

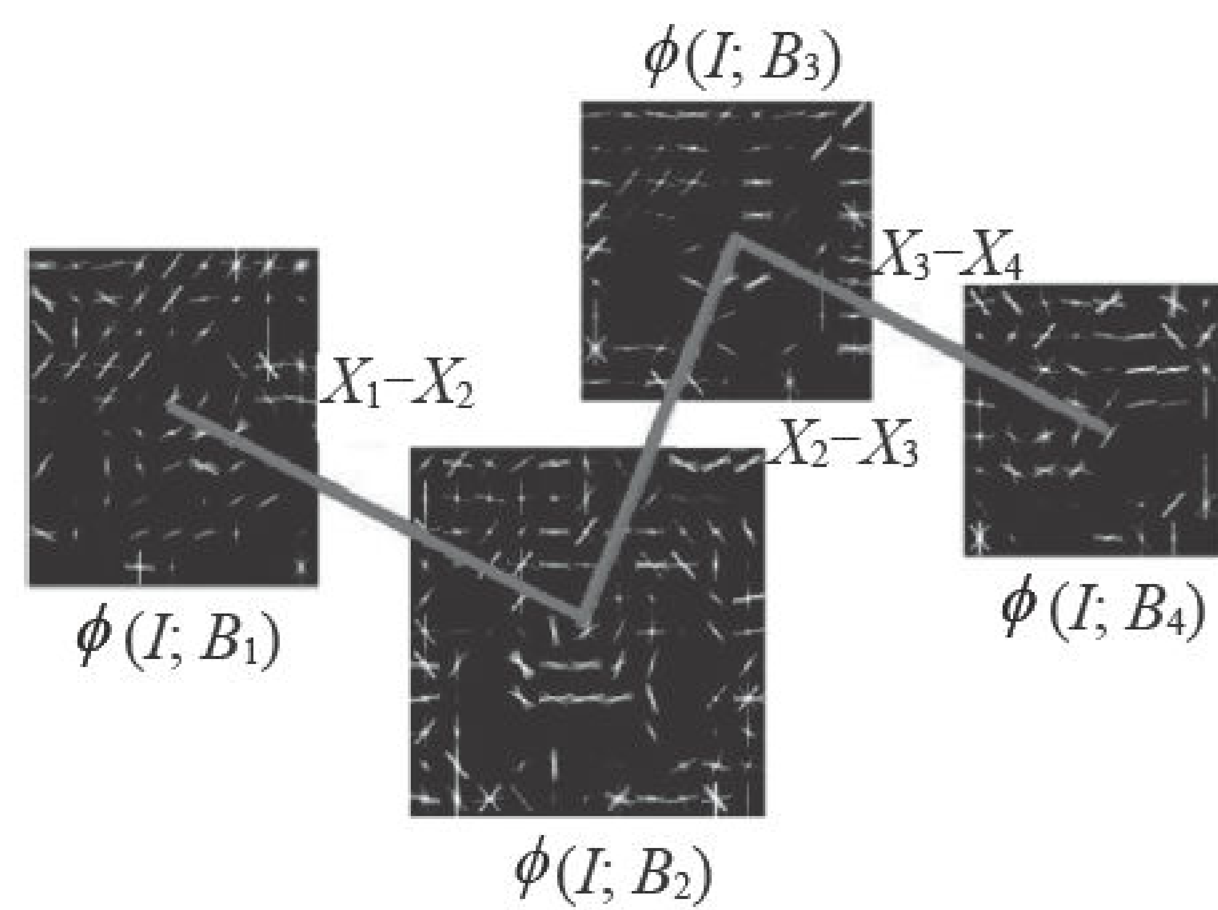


图8-9 提取特征

接下来，将待检测图像块的HOG特征输入线性SVM进行训练，输出相应的HOG特征模板，即分类器权值，再经过与SVM训练模板卷积操作，即可获得目标的参数化外观模型。另外，根据最小生成树模型确定目标间边的连接关系，即目标间空间位置关系的构建。最后，多目标跟踪模型建成，如图8-10所示。

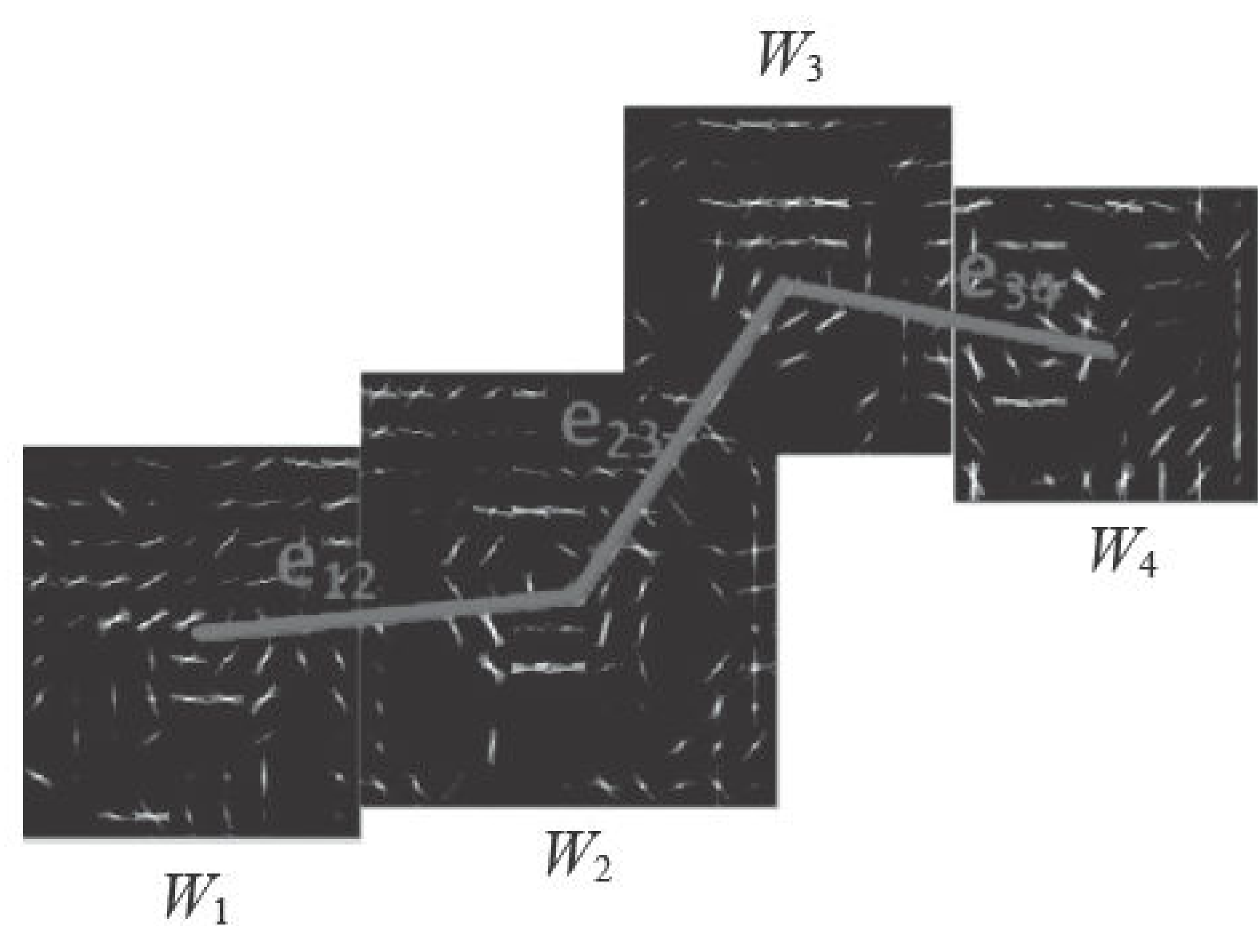


图8-10 完成模型构建

从图8-10中可以看出，利用HOG特征模板可以突出目标的稳定特征，弱化不稳定特征，使目标外观模型更加准确。分类器输出就是该定位区域的图像信息与该目标部位真实图像块相匹配。

第9章

SVM在线学习的模型参数更新



9.1 在线学习的跟踪算法

传统的跟踪算法可看作一类搜寻匹配问题，利用目标模型在视频中不断搜索，找到相似度最高的区域。该目标模型由大量样本数据离线建立，还可以根据更新公式进行定时更新。目前，该算法主要用于跟踪一些如车辆等变化不明显的刚性目标，但对于行人、动物等外形多变的柔性目标，则不能实现目标的鲁棒性跟踪。采用该方法容易发生跟踪漂移，导致跟踪失败的原因大致有两点：首先，在离线状态下，要想建立一个完全包含目标所有状态的模型几乎是不可能的；其次，由于柔性目标的外形变化随机，毫无规律可循，简单按照更新公式更新模型只能在一定程度上解决柔性目标的外形多变问题，容易出现跟踪漂移现象。另外，保证该离线分类器正常工作的前提是需要人工标记大量的训练样本，这是非常费时费力的。例如，Zhang[213]等利用该方法做人脸识别实验，标记了近11 000个正样本和100 000个负样本图片，而且样本训练工作完成后，离线分类器的各种参数就已被确定，使得该分类器很难推广到任意的目标类型。

在线分类器是指初始化后具有一定的分类能力的分类器，能应用到其后图像序列的分类，并将判别结果继续作为训练样本进行实时更新。[214]在线更新分类器最显著的特点是具有较强的自适应能力，可以实时适应目标外表的变化。在线学习跟踪方法的本质是检测视频中未标记的数据并对其进行分类。[215]它能把未标记的数据进行类别区分，并把正样本存入样本数据库中。在跟踪过程中，随着时间的推移，数据库中的目标模型逐渐得到完善，同时数据库中正样本不足的问题也得以解决。这种新思路完全区别于传统方法的套路，不再设法建立复杂的离线模型与更新规则来描述目标及目标的变化，而是在视频图像序列中，根据目标检测模型和一定的空间约束条件来搜寻目标位置。显然，在线学习跟踪方法将目标的跟踪与检测看作一个整体，这种思想为目标跟踪的研究开辟了新方向。

9.2 SVM在线学习的跟踪框架

本书将跟踪问题看作一种特殊的目标检测问题，根据目标检测模型和一定的空间约束条件，在视频中重新搜索并定位目标的位置。我们利用跟踪目标在先前帧中的配置作为正

样本来更新我们的模型，在一个在线的结构化的SVM框架中训练所有目标的外观模型和这些目标间的结构约束。基本框架如图9-1所示，分类器通过这种在线学习和训练的方式来完成目标跟踪。

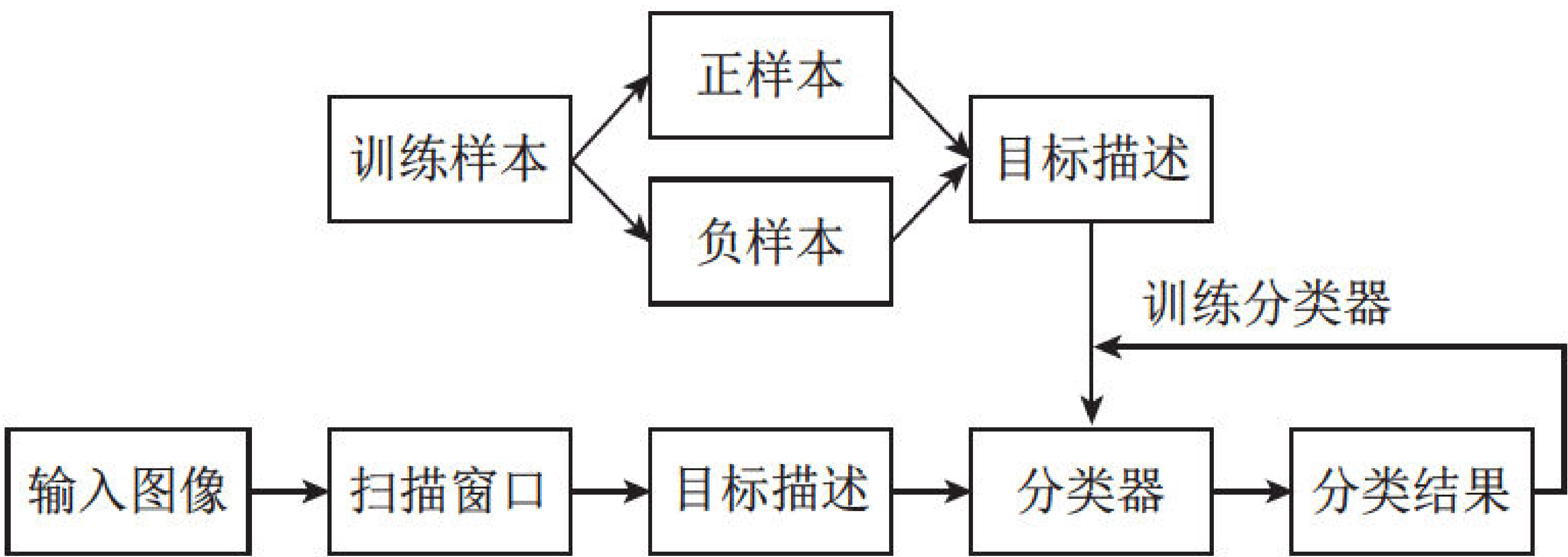


图9-1　在线学习跟踪的基本框架

实现SVM在线跟踪算法的具体细节如下所述。

1. 样本初始化

第一帧数据往往是通过人为标记的，也可以利用目标检测算法，如帧间差分、背景差分等算法来实现。本书首先采用人工标记的方法确定初始帧中目标矩形框的坐标位置和区域大小，从而得到准确的目标信息。对于大部分目标来说，可利用一帧图像得到大量的正负样本，提取每个样本的HOG特征，建立初始化训练样本集。这里，在第一帧图片上根据计算包含目标区域的图像块特征来获得表示目标的正样本，同样，在除了目标矩形框以外的任意位置，提取含背景区域的等大图像块特征作为负样本。把这些训练样本集作为SVM分类器的输入进行训练，从而使分类器在一定程度上具有分类能力。

2. 目标定位

提取待检测样本输入分类器，根据输出结果预测在新一帧中目标的位置定位，完成该帧的目标跟踪。该算法能从已有的跟踪模型基础上利用分类器输出未标定帧的置信图，由置信度峰值确定这一帧目标的位置。整个视频目标跟踪可看作通过机器学习、模式识别等，完成一个训练与识别的交替循环过程。

3. 更新分类器

目标确定的位置配置被用于更新分类器，小的定位误差可能会随着时间的推移而传播，导致跟踪漂移。因此，对于在线学习的跟踪算法，要根据每帧中对目标跟踪结果的置信度，判断是否把当前该帧的目标位置配置送入训练样本特征库中，作为正样本用于更新分类器。当我们的目标是遮挡的，我们不想更新该目标的外观模型。为了避免对外观模型

的错误更新，我们设定固定阈值，只有当目标的置信度大于该阈值时，我们才更新检测框对应的分类器权值。否则，分类器将跳过更新步骤，转入下一帧的跟踪。

9.3 结构化的SVM学习

为了适应在跟踪过程中目标外观的剧烈变化和目标间相互位置关系的变化、表示模型需要进行实时更新以此来避免跟踪框的偏移情况，本书将跟踪图像序列中每帧多目标的位置配置输出作为正样本，用来学习外观模型参数ω_i和目标间的相互关系参数e_{ij}。我们将对模型参数$\Theta=\{\omega_1, \dots, \omega_{|V|}, e_1, \dots, e_{|E|}\}$进行结构化学习(Structured SVM)。

9.3.1 结构化的SVM

一般的机器学习方法在建模时，通常需要寻找一个映射函数：$f: X \to Y$，输入X映射为Y。这里，输入X的形式是多种多样的，在一般的机器学习方法中，输出的Y通常是一个数值或者一个0或者1的标签，比如分类、回归、分布密度估计等方法。但是，在结构化学习方法中，Y不再局限于作为一个数值，而是一个复杂的结构对象。比如，输入一幅图像得到的是一个带标签的序列或者是一个稀疏的树桩结构。因此，结构化的学习方法能够直接解决实际问题，为直接得到具体问题中的输出提供了一种框架，在理论上，可以把任何一种输出都认为是一种结构。

在结构化的学习中，训练过程就是根据训练集(x_1, y_1)，(x_2, y_2)，…，(x_n, y_n)训练一个模型。这里，模型通常是一个参数化模型，学习的过程就是一个确定参数值的过程。结构化学习包括很多方法，其中比较重要的一类方法就是基于概率图模型的方法，尤其是无向图模型，将结构化学习的问题转化为图模型中的基本问题。[216]另一类方法就是Large Margin Method for Structured Learning，也就是Structured SVM，是对SVM方法的一种扩展。目前，Structured SVM已经具备了比较成熟的算法框架。[217]

在Structured SVM中，主要问题是要寻找一个判别函数：$F: X \times Y \to R$；给定一个输入值，选择能使判别函数F最大化的Y值作为输出，即

$$f(x;\ w)=\arg\max_{y\in Y}F(x,\ y;\ w) \tag{9-1}$$

这里，w表示一个参数向量；可把F理解为一个度量$(x,\ y)$兼容程度的函数，表示为$F(x,\ y;\ w)=w\cdot\varphi(x,\ y)$，其中，$\varphi(x,\ y)$是输入对输出的特征表示，其形式取决于具体问题。

为了量化预测的准确性，定义了一个损失函数$\Delta: Y\times Y\rightarrow R$，给定一组训练集$S$，$f(x;\ w)$的性能可以通过经验损失函数来度量。不同的$f(x;\ w)$对应不同的经验损失，经验损失越小越好。

经验损失为0意味着得到一个完全符合训练样本要求的判定函数f，即$\forall(x_i,\ y_i)$，$f(x_i)=y_i$，但一般无法得到满足约束的可行解。因此，在借鉴SVM算法思想的基础上加入松弛变量ξ_i，使模型不必完全拟合训练集中的样本，从而得到下面的优化问题，即

$$\begin{aligned}&\min_{w,\ \xi}\frac{1}{2}\|w\|^2+\frac{C}{n}\sum_{i=1}^{n}\xi^2\\&s.t.\forall i,\ \forall y\in Y\setminus y_i:\ \langle w,\ \delta\varphi_i(y)\rangle\geqslant 1-\xi_i\\&\text{其中，}\delta\varphi_i(y)=\varphi(x_i,\ y_i)-\varphi_0(x_i,\ y)\end{aligned} \tag{9-2}$$

这里可以将ξ_i理解为对不满足约束条件的样本y的惩罚。但是，这里对于两个不满足约束条件的y_1和y_2，如果$\Delta(y_1,\ y_i)>\Delta(y_2,\ y_i)$，那么将给予更大的惩罚。因此，在惩罚时应该考虑损失函数的影响，可以进一步得到以下优化问题，即

$$\begin{aligned}&\min_{w,\ \xi}\frac{1}{2}\|w\|^2+\frac{C}{n}\sum_{i=1}^{n}\xi^2\\&s.t.\forall i,\ \forall y\in Y\setminus y_i:\ \langle w,\ \delta\varphi_i(y)\rangle\geqslant\Delta(y_i,\ y)-\xi_i\end{aligned} \tag{9-3}$$

通过上面的分析，可以将求解函数f转化为SVM的优化问题，通过解决这个问题，可以得到函数f。由于上述优化问题的约束条件太多，直接使用SVM进行求解是不可行的，需要使用特定的方法才能求解。

9.3.2　结构化学习

若一帧视频图像I对应的最佳位置输出为C，则学习参数为Θ，使最优值满足Structured SVM的损失$\ell(\Theta,\ C)$最小，则

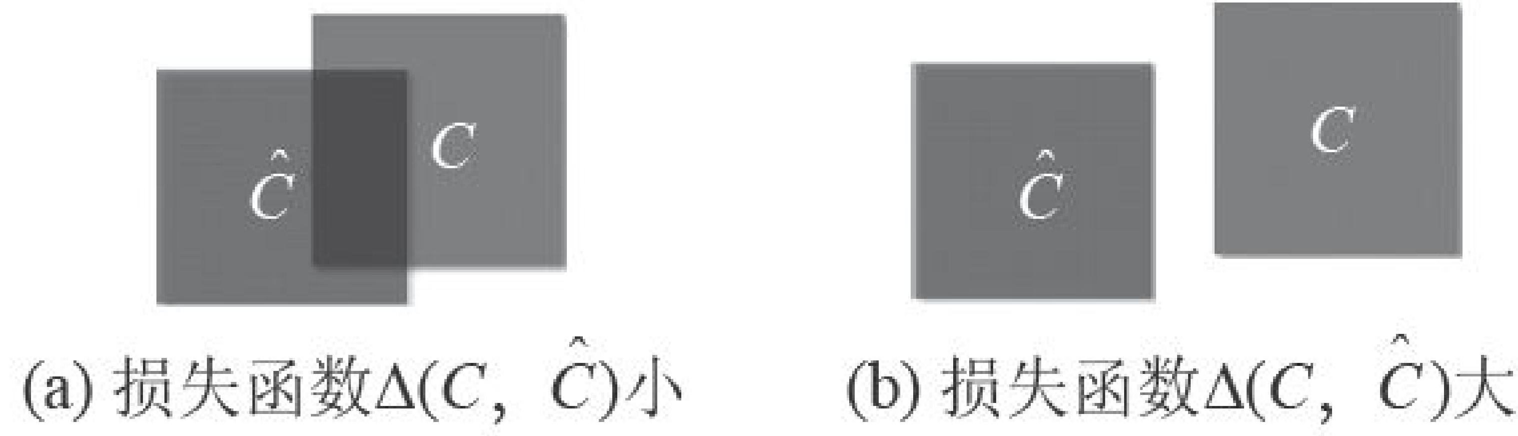
$$\ell(\Theta,\ C)=\max_{\hat{C}}\left[s(\hat{C},\ \Theta)-s(C,\ \Theta)+\Delta(C,\ C)\right] \tag{9-4}$$

其中，C为多个目标的真实位置配置；$\hat{C}$为候选位置配置；$\Delta\left(C，\hat{C}\right)$为损失函数，用来衡量当前分类器对位置配置预测的准确程度。$\Delta\left(C，\hat{C}\right)$满足以下条件：当且仅当$\hat{C}=C$时，$\Delta\left(C，\hat{C}\right)=0$；$\Delta\left(C，\hat{C}\right)$是非负的，存在$\forall\hat{C}\neq C$，有$\Delta\left(C，\hat{C}\right)>0$；存在一个常数$M$有上界，对于$\forall C$有$\max_{\hat{C}}\Delta\left(C，\hat{C}\right)<\mathrm{M}$。在我们的跟踪器中，根据$C$和$\hat{C}$之间的重叠度来定义损失函数，即

$$\Delta\left(C，\hat{C}\right)=\sum_{i\in V}\left(1-\frac{B_i\cap\hat{B}_i}{B_i\cup\hat{B}}\right) \tag{9-5}$$

其中，两个边界框的并集与交集在像素中测量。

损失函数的形象化表示，如图9-2所示。



(a) 损失函数$\Delta(C，\hat{C})$小　　(b) 损失函数$\Delta(C，\hat{C})$大

图9-2　损失函数示意图

结构化的SVM损失$\ell\left(\Theta, C\right)$不包括二次项，但它是仿射函数的一个集合的最大值。因此，公式(9-4)中的$\ell\left(\Theta, C\right)$可以看作参数$\Theta$的一个凸函数。故在其方向上使用一个梯度步骤的搜索方向P来学习我们的模型参数Θ。Structured SVM的梯度可以表示为

$$P=\nabla_{\Theta}s\left(C^*，\Theta\right)-\nabla_{\Theta}s\left(C，\Theta\right) \tag{9-6}$$

其中负配置C^*给出

$$C^*=\arg\max_{\hat{C}}\left[s\left(\hat{C}，\Theta\right)+\Delta\left(C，C\right)\right] \tag{9-7}$$

为了从实质上减小损失，同时确保参数更新的变化不太大，我们使用被动-主动(Passive Aggressive，PA)算法来完成参数更新。被动-主动算法[218]是一种感知器学习算法，在修正权值中加入主动参数，即$w_{t□1}=w_i+\tau_i y_i$。在原目标函数$\begin{matrix} w_{t+1}=\arg\min\limits_{w\in R^n}\frac{1}{2}\left\|w-w_t\right\| \\ s.t.\ \ l\left[w;\left(x_t，y_t\right)\right]=0 \end{matrix}$的基础上，针对样本的线性可分与不可分的情况，分别对目标函数进行修正。加入松弛变量ξ得到PA□I：$\begin{matrix} w_{t+1}=\arg\min\limits_{w\in R^n}\frac{1}{2}\left\|w-w_t\right\|+C\xi \\ s.t.\ \ l\left[w;\left(x_t，y_t\right)\right]\leqslant\xi，\xi\geqslant 0 \end{matrix}$，加入松弛变量$\xi^2$得到PA□II：$\begin{matrix} w_{t+1}=\arg\min\limits_{w\in R^n}\frac{1}{2}\left\|w-w_t\right\|+C\xi^2 \\ s.t.\ \ l\left[w;\left(x_t，y_t\right)\right]\leqslant\xi \end{matrix}$。

目标函数将依次得到权值修正时所对应的三个主动参数[219]：$\tau_t = \frac{l_t}{\|x_t\|^2}$，$\tau_t = \min\left\{C, \frac{l_t}{\|x_t\|^2}\right\}$，$\tau_t = \frac{l_t}{\|x_t\|^2 + \frac{1}{2C}}$。

在本书中，结构化的预测器是用基于前帧检测的在线方式来训练的。在每一帧用一个简单的随机子梯度下降算法后，我们更新目标跟踪器。每个参数更新的步长由被动-主动算法决定。可得

$$\Theta \leftarrow \Theta - \frac{\ell(\Theta, C)}{\|P\|^2 + \frac{1}{2k}} P \tag{9-8}$$

其中，P为学习过程中的搜索方向；$k \in (0, +\infty)$为超参数，表示控制参数在更新过程中的“主动”程度。实验中，固定$k=1$。

9.4　实验结果与分析

实验借助LibSVM函数库、Matlab等有效的工具，通过编写程序完成对算法有效性的检验。程序中主要包括HOG特征的提取、MST模型的构建、输入图像的检测等模块。本节共进行了4组跟踪实验，均选取具有挑战性影响因素的视频段，视频信息如表9-1所示。这些视频涉及遮挡、变形、相似外观以及背景和光照多变等在目标跟踪领域常见的难处理的因素。实验采用的实验硬件是Pentium(R) Dual-Core CPU E5300，主频2.60GHz，内存2.00GB。实验在Matlab R2011a上进行。

表9-1　视频序列信息

视频序列	名称	总帧数	难度
1	Carchase	350	背影同化、遮挡
2	Red Flowers	2249	相似外观
3	Hunting	1805	姿态多样性
4	Shaking	344	光照变化

本书基于置信图显示分类结果。置信图由灰度值显示待测图像中各像素对目标的置信度，表达利用某显著性特征信息区分目标与背景的预测效果。[220]本书算法能从已有的跟踪模型基础上利用分类器输出未标定帧的置信图，由置信度峰值确定这一帧目标的位置。同时，依据每次跟踪结果的置信程度决定是否利用当前帧在线更新正负样本和目标间的位置关系。这里我们设定阈值为0.4，假如目标置信度比这个阈值小，则跳过更新步骤转入下一帧的跟踪。

首先，定性地分析视频跟踪的效果。如图9-3所示，在车辆追逐视频中，背景在不断地发生变化；从图像的第200帧开始，两辆车会通过一座大桥，在这个过程中目标被完全遮挡，但由于本书算法的跟踪器可根据一辆车的位置估计另一辆车的位置，跟踪器会很快追踪到目标位置。

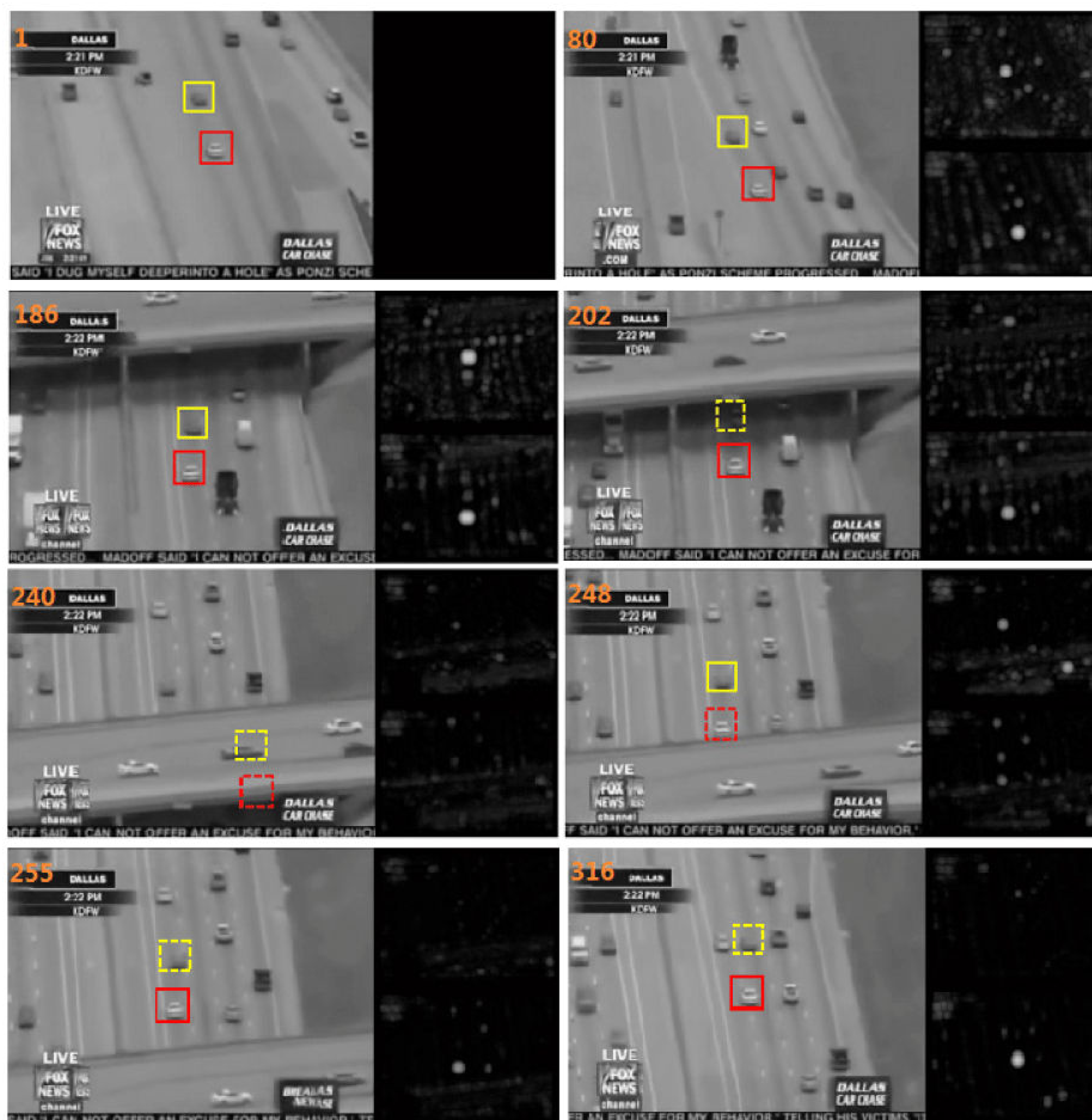


图9-3　跟踪结果(帧号依次为1，80，186，202，240，248，255，316)

第二组视频跟踪几朵外形相似的花朵，由于风力的影响，存在方向移动和外形改变，有时发生相互间部分遮挡。本书算法的跟踪器可以利用结构约束来区分具有相似外形的不同花朵，防止跟踪器在目标间发生变换，跟踪结果如图9-4所示。第一张图片是视频的第一帧中目标的真实标注。每幅图片的右侧显示当前帧的置信图，由此可以看出在整个视频长度(2249帧)中能很好地跟踪。

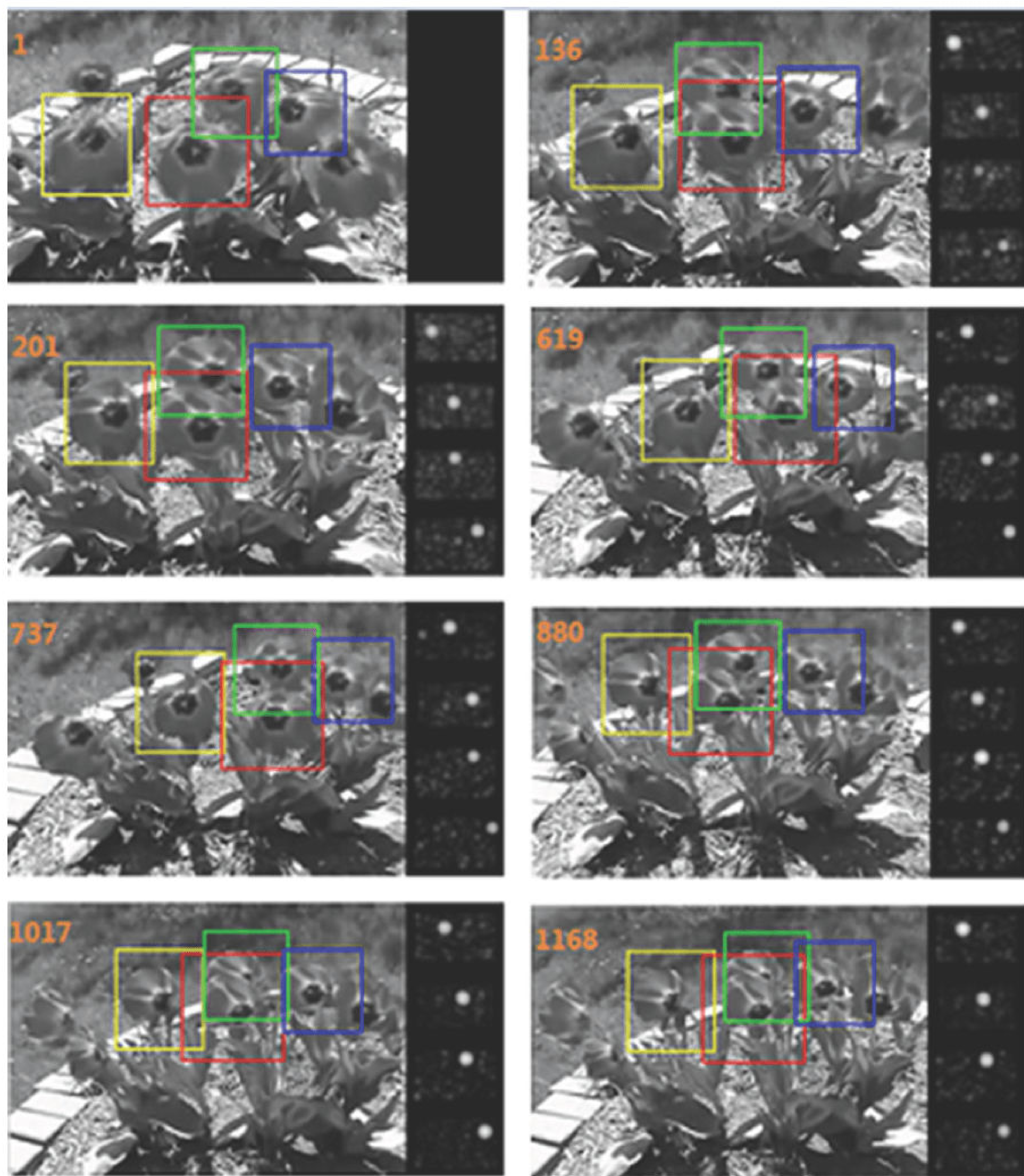


图9-4　跟踪结果(帧号依次为1，136，201，619，737，880，1017，1168)

在第三组视频中，猎豹和小羚羊的外观和相对位置随着时间的推移发生显著改变，这对跟踪的鲁棒性具有很大的挑战。由于两个动物在相同方向上移动，故跟踪器通过空间的结构约束实时更新两个目标间的空间距离，因此依然可以保持鲁棒性，进行有效的跟踪，

跟踪结果如图9-5所示。

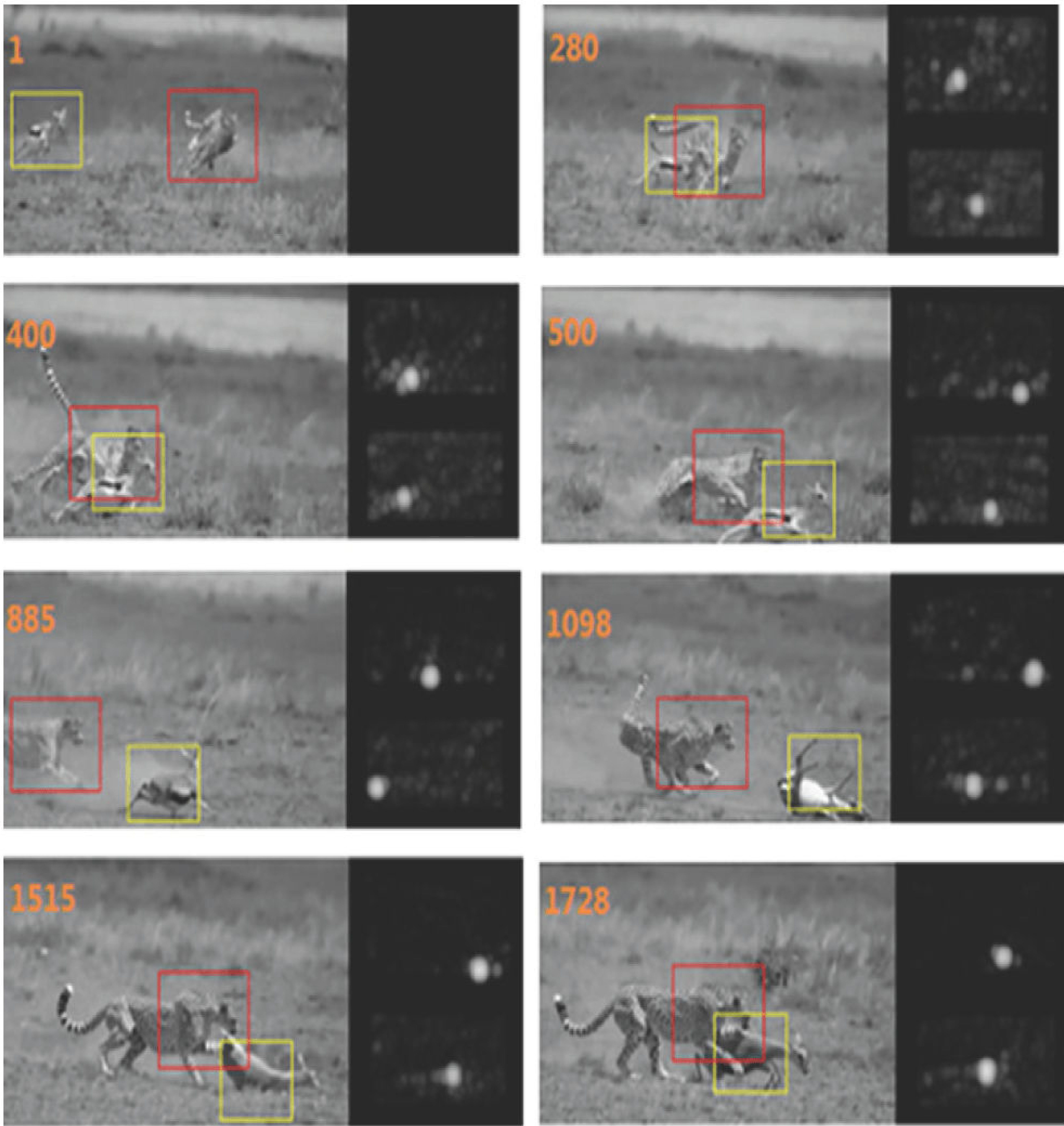


图9-5　跟踪结果(帧号依次为1，280，400，500，885，1098，1515，1728)

第四组视频跟踪器要同时跟踪三个外观变化不大的目标，但是，由于光线多变，会引起背景发生很大变化；另外，光线昏暗使目标和背景的区分度较低，这都给跟踪工作带来很大的挑战。本书算法的跟踪器利用结构约束可有效完成对多目标的跟踪任务，如图9-6所示。

接下来，我们统计两个指标来定量地评价跟踪器：平均定位误差和正确检测率，更客观、直接地显示该算法对多目标场景的跟踪性能。平均定位误差(Average Location Errors，ALE)是通过计算预测边界框的中心位置到人工标定的目标真实边界框的中心位置之间的平均欧氏距离得到的，误差越小越好。正确检测率(Correct Detection Rate，CDR)是指预测

边界框和标准边界框之间的重叠率不小于50%的帧数比率，比率越大越好。本书在同一实验环境下，与不含目标间结构约束的传统HOG+SVM跟踪器做比较，来评价我们的跟踪器的性能。在对整个图像序列的跟踪过程中，依次统计了两种算法下所有帧跟踪结果的ALE和CDR的平均值，如表9-2和表9-3所示。

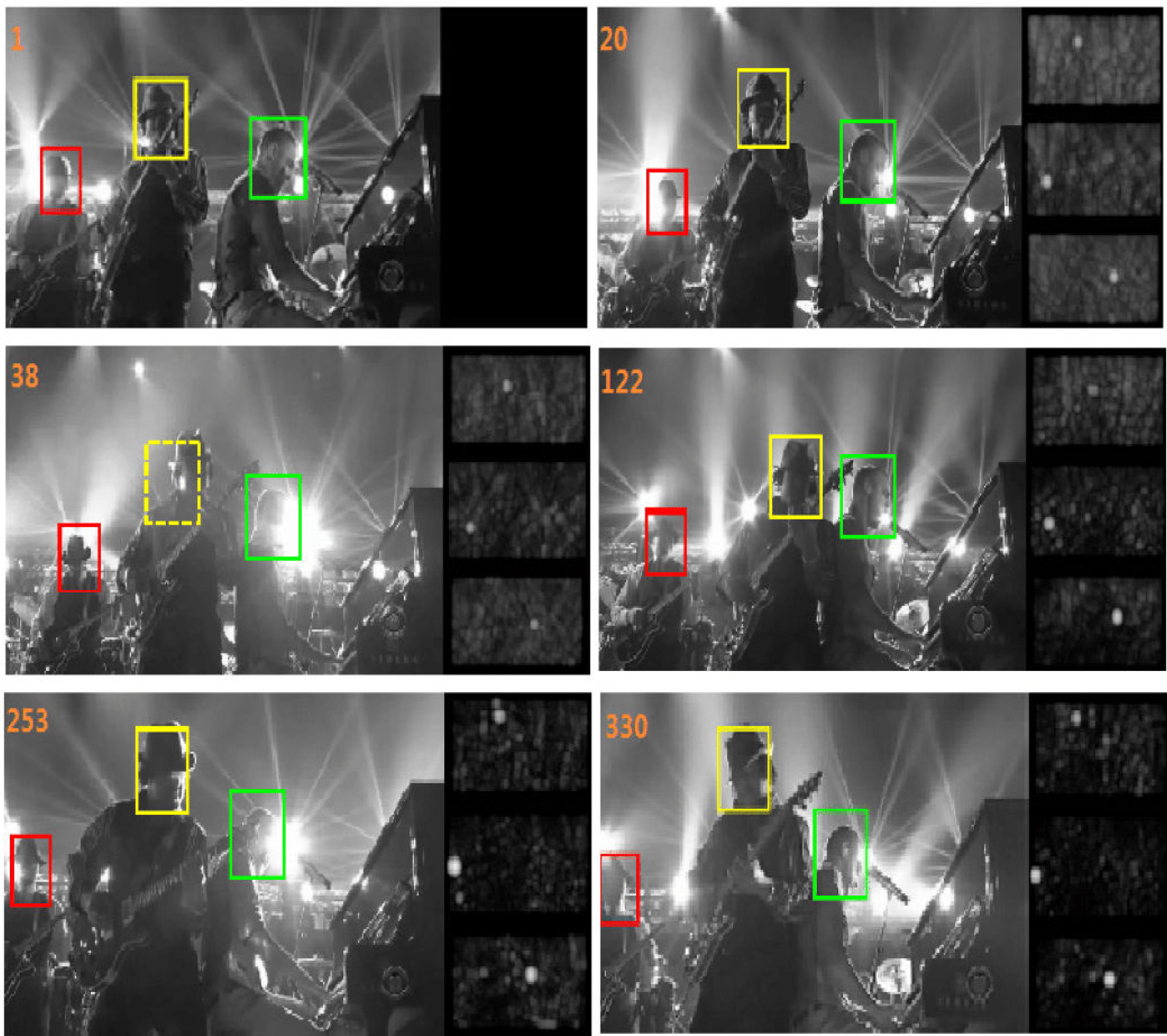


图9-6　跟踪结果(帧号依次为1，20，38，122，253，330)

表9-2　平均定位误差

视频名称	本书算法	HOG+SVM算法
Car Chase	7.5	24.8
Red Flowers	9.5	50.6
Hunting	19.4	171.7
Shaking	7.7	58.3

表9-3　正确检测率

视频名称	本书算法	HOG+SVM算法
Car Chase	0.91	0.78
Red Flowers	0.99	0.38
Hunting	0.87	0.07
Shaking	0.97	0.47

由表9-2和表9-3可以看出，本书算法的平均定位误差值均小于Dala-Triggs检测器(它采用HOG特征描述图像块，用SVM预测目标的存在，通过同时对多个目标进行单目标跟踪来最终实现多目标跟踪)的平均定位误差，而检测正确率均大于Dala-Triggs检测器。故本书改进的跟踪器能实现更加准确的跟踪，其在多目标场景中的性能得到提高。

第10章
总结与展望



10.1 工作总结

我们对无人机航空遥感图像的拼接计算，进行了以下几方面的研究工作。

(1) 根据无人机遥感图像成像的内、外方位元素，采用直角空间变换及二次线性插补方法，实现了遥感图像校正。根据获取的无人机飞行状态参数，实现了连续两幅无人机航空遥感快视图像之间重叠区域的图像范围计算。

(2) 提出分布存储环境下的并行几何校正算法，每个处理器通过计算本地输入子图像在目标图像中的范围，确定其需要进行重采样计算的区域，使计算过程中所需的数据均为本地数据，很好地解决了数据局部性问题。利用首尾相连的闭线段近似表示理想的输出图像块边界，详细讨论了局部输出区域的计算方法，并采用一种新的存储结构用于保存校正后的输出图像块信息。

(3) 基于图像数据总体分布的统计，分析了飞行试验图像的成像质量。基于人眼亮度视觉特性曲线，结合小波变换和Curvelet变换的特点，提出一种新的图像增强方法，实现了无人机遥感序列图像的自适应增强处理。

(4) 基于匹配特征提取方法，提出了一种将小波多分辨率分析特性和Canny算法相结合的检测图像大边缘的方法，能够较好地从图像中提取出比较完整的大边缘，而忽略一些小的边缘。该边缘提取方法为图像配准提供了一个良好的匹配特征。

(5) 将图像匹配分为粗匹配和精匹配两个步骤。粗匹配计算中，首先确定待拼接的两幅图像之间的大致重叠区域，利用小波变换与Canny算法提取图像边缘。采用区域匹配方法求得两幅图像的匹配点。精匹配计算中，在更小的搜索区域，利用最小二乘法，计算得到待拼接图像之间的最佳相对位置关系。根据匹配结果，实现两幅图像的拼接。

(6) 基于人眼的颜色视觉特性分析，提出了一种具有抗亮度干扰能力的彩色图像色差度量方法。利用颜色相似性分析，并引入协方差矩阵计算，给出了彩色图像特征模板的提取方法。利用最小二乘法，建立了两幅彩色图像之间的亮度变换函数曲线，实现了基于基准图像亮度分布的伽马校正。

(7) 目标的特征描述是跟踪算法中的重要步骤，直接影响最终的跟踪效果，决定整个跟踪系统的性能。基于分类学习的目标跟踪一般不选择较直观的整体特征，而偏向于由基

本特征进行局部特征统计，使特征描述子更易于区分且计算量较小。选择目前成功应用于目标检测的HOG特征作为目标特征。

(8) 传统的离线分类器是需要人为标记大量样本的训练分类器，当待测样本与训练样本有较大差异时，无法及时适应目标的外观变化，且难以推广到其他目标类型，分类器应用具有局限性。基于检测的跟踪算法，在跟踪过程中，跟踪器根据上一帧的跟踪结果信息在线更新分类器，然后再对当前帧检测，以获取当前运动目标的准确预测位置，有效抑制跟踪漂移。首先，构造目标的外观模型。在初始帧中随机选择大量含目标区域以及含背景区域的等大图像块，分别提取HOG特征，标记为正负样本。然后，将这些样本特征集合作为训练样本输入SVM分类器，即可训练得到目标准确的参数化外观模型。由于HOG特征具有几何不变的特性，经过SVM训练后得到的目标外观模型将更加准确。本书所提出的遥感图像处理算法，实现了仿真程序设计，验证了算法的可行性，完成无人机航空遥感图像动态拼接的软件设计工作。

10.2　未来工作的展望

遥感图像拼接是一个实用但复杂的数据处理过程，本书所提出的算法在以下几方面需要做进一步研究。

(1) 图像校正时，只是针对所摄图像在水平方向上的旋转进行了校正，而忽略了俯仰和横滚的旋转，使图像校正难以达到更高的精度。

(2) 多帧对齐优化技术，包括基于平面景物模型的对齐优化技术和基于旋转运动模型的对齐优化技术，它们是提高图像Mosaic质量和自动计算相机非线性畸变参数的基础，我们将建立和推导旋转运动模型，由于多帧对齐优化的运算量较大，在实际应用中可能是无法承受的，因此寻找快速、减少运算代价的算法是未来的研究内容。

(3) 自动计算相机的非线性畸变参数：利用光强度常量的约束建立并推导带有非线性畸变参数的多帧优化目标函数，分别使用基于平面景物模型的对齐优化技术和基于旋转运动模型的对齐优化技术来自动计算相机的非线性畸变参数。

(4) 建立动态图像Mosaic，即建立基于图像的动态虚拟环境的基础。我们主要研究图

像序列间的对齐，针对机载相机拍摄的两个视频序列，使用所有的时空信息来估计空间帧的对齐和两个序列的同步。动态环境(包括移动物体和光源变化)是对齐运算的有利线索，我们将研究瞬间直接对齐技术而不需要事先估计点和帧的对应关系，它把时间和空间线索合成一个统一的框架来建立优化目标函数。

以上不足尚需进一步完善，以提高图像拼接的精确性和快速性。

参考文献

[1] 钟力，胡晓峰. 重叠图像拼接算法[J]. 中国图象图形学报，1998，3(5)：367-370.

[2] 谢凤英，姜志国. 基于互相关的显微医学图像配准[J]. 中国体视学与图像分析，2001，6(3)：175-178.

[3] 殷世民，刘上乾，吴志鹏，等. 一种基于特征区域分割的图像拼接算法[J]. 西安电子科技大学学报，2002，29(6)：768-771.

[4] Kim D H，Yoon Y I，Choi J S. An Efficient Method to Build Panoramic Image Mosaics[J]. *Pattern Recognition Letters*，2003，24(14)：2421-2429.

[5] 王建忠，肖绍良. 图像镶嵌及其边界处理[J]. 模式识别与人工智能，1993，6(3)：189-195.

[6] 徐正光，田清，张利欣. 图像拼接方法探讨[J]. 微计算机信息，2006，22(10)：255-257.

[7] 李从利，韩裕生，冷晓艳，薛模根. 序列图像拼接过程中误匹配的消除[J]. 计算机应用与软件，2006，23(11)：93-97.

[8] 朱述龙，钱曾波. 遥感图像镶嵌时拼接缝的消除方法[J]. 遥感学报，2002，6(3)：183-187.

[9] 曹清. 大型图像数据库中的色调调整方法[J]. 武汉大学学报：信息科学版，2003，18(3)：311-314.

[10] 杜志军，贾平. 灰度投影在块匹配算法中的应用[J]. 微计算机信息，2007，23(2)：269-272.

[11] 徐玮，张茂军，吴玲达. 虚拟实景空间的颜色校正方法研究[J]. 小型微型计算机系

统，2001，22(2)：165-167.

[12] 葛仕明，程义民，曾丹，李杰. 基于梯度场整体变分模型的无缝图像处理方法[J]. 中国科学院研究生院学报，2006，23(5)：665-669.

[13] 赵卿松. 序列图像拼接中的交互调整方法[J].计算机工程，2007，33(4)：201-203.

[14] 周廷刚. 遥感彩色图像镶嵌拼接缝的消除方法研究[J]. 计算机工程与应用，2004，36：84-86.

[15] 沈项军，李峰，陈金华. 一种基于特征跟踪的彩色序列图像拼接方法[J]. 微型电脑应用，2003，19(3)：39-41.

[16] 严大勤，孙鑫. 一种基于区域匹配的图像拼接算法[J]. 仪器仪表学报，2006，27(6)：749-753.

[17] 高富强，张帆. 一种快速彩色图像匹配算法[J]. 计算机应用，2005，25(11)：2604-2605，2611.

[18] 冷晓艳. 空中侦察图像连续拼接中的平滑算法研究[J]. 计算机工程与科学，2007，29(2)：52-55.

[19] 王伟，陆佩忠. 数字图像拼接技术[J]. 小型微型计算机系统，2006，27(7)：1347-1352.

[20] 李胜睿，李翠华. 基于梯度滤波的彩色图像快速拼接与实现技术[J]. 厦门大学学报：自然科学版，2003，43(1)：29-34.

[21] 王平江，甄恒洲. 特征稀少图像的拼接方法[J]. 华中科技大学学报：自然科学版，2006，34(8)：13-16.

[22] 赵晓敏. 一种基于图像投影特征的射线图像拼接方法[J]. 光学技术，2006，32(6)：929-934.

[23] Hisausw.*CIE Industrial color difference evaluation*[M]. Vienna：CIE Publication，1995.

[24] Berns R S，Alman D H，Reniff L et al. Visual determination of supra threshold color difference tolerances using profit analysis[J]. *Color Research and Application*，1991，16(5)：297-299.

[25] Levkowitz H，Herman G T. Color scales for image data[J]. *IEEE Computer Graphics and Application*，2006，12(1)：72-75.

[26] Cyheuew G. *Studio encoding parameters of digital television for standard 4：3 and*

wide-screen 16：9 aspect ratios[M]. Geneva：ITU-R，Recommendation BT. 601-5，1994.

[27] Ohta Y，Kanade T，Sakai T. Color information for region segmentation[J]. *Computer Graphics and Image Processing*，1980，13(2)：222.

[28] Pratt W K. *Digital image processing*[M]. 2nd ed. New York：John Wiley，1991.

[29] Foley J D，van Dam A. *Fundamentals of Interactive Computer Graphics*[M]. MA：Addison-Wesley，1982.

[30] Ledley S，Buas M，Golab T. Fundamentals of true-color image processing[C]，Proceedings of 10th International Conference on Pattern Recognition. Atlantic City，1990，791-795.

[31] Gonzalez R C，Woods R E. *Digital image processing*[M]. 3rd ed. Reading，MA：Addison Wesley，1992.

[32] Bajon J，Cattoen M，Kim S D. Real-time colorimetric transformations used in robot vision[C]. Proceedings of MICAD，French，1985，76-86.

[33] Machuca R，Phillips K. Application of vector fields to image processing[J]. *IEEE Trans PAMI*，1983，(5)：316-329.

[34] Pitas I，Venetsanopoulos A N. *Nonlinear digital filters principles and applications*[M]. New York：Kluwer Academic Publishers，1990：129-138.

[35] Astola J，Haavisto P，Neuvo Y. Vector median filter[C]. Proc. of IEEE Int. Conf. on Image Processing，1990，78：678-689.

[36] Trahanias P E，Venetsanopoulos A N. Vector directional filters proposed class of multichannel image processing filters[J]. *IEEE Trans of Image Processing*，1993，(2)：528-534.

[37] Karakos D G，Tranias P E. Combining vector median and vector directional filters：the directional-distance filters[C]. Proc of IEEE Int. Conf on Image Processing 95，1995，1：171-174.

[38] Plataniotis K N，Androutsos D，Venetsanopoulos A N. Content-based color image filters[J]. *Electronics Letters*，1997，(33)：203-212.

[39] Buchowicz A，Pitas I. Multidimensional distance filters[C]. Proc of IEEE Int Conf on Image Processing 94，1994，2：575-579.

[40] 郁伯康，郁梅. 彩色图像滤波的新方法[J]. 电子科技大学学报，1999，28(1)：14-19.

[41] Ralph Sucher. A recursive nonlinear filter for removal of impulse noise[J]. *Proc. of the IEEE Int. Conf. on Image Processing (ICIP'95)*，1995：183-186.

[42] 蒋刚毅，郁梅，郁伯康.自适应彩色图像滤波新方法[J]. 科学通报，1999，9：1937-1942.

[43] 李永峰，丁润涛，国澄明. 一种彩色图像滤波的改进矢量中值算法[J]. 数据采集与处理，2003，18(3)：361-364.

[44] 郁梅，郁伯康，蒋刚毅，郑义. 基于噪声检测的彩色图像脉冲噪声滤波[J]. 计算机工程与应用，2000(3)：55-57.

[45] 王文惠，楼生强，万建伟，周良柱. 一种彩色图像的非线性自适应滤波算法[J]. 计算机工程与科学，2001，23(4)：18-20.

[46] 郁伯康，郁梅. 彩色图像脉冲噪声的自适应矢量滤波[J]. 电子科技大学学报，2000，29(2)：131-135.

[47] 章毓晋. 图像分割[M]. 北京：科学出版社，2001.

[48] Pal N R，Pal S K. A review on image segmentation techniques[J]. *Pattern Recognition*，1993，(26)：1277-1294.

[49] Uchiyama T，Arbib M. A colour image segmentation using competitive learning[J]. *IEEE-PAMI*，1994，(16)：1197-1206.

[50] Lim Y W，Lee S U. On the color image segmentation using algorithm based on the thresholding and fuzzy C-means techniques[J]. *Pattern Recognition*，1990，23(9)：935-952.

[51] 柏子游，张勇，虞烈. 一种彩色图像的色彩分割方法[J]. 模式识别与人工智能，1999，12(2)：241-244.

[52] 郭国栋，马颂德. 彩色图像分割[J]. 中国图象图形学报，1998，3(11)：918-921.

[53] Ohlander R，Price K，Reddy D R. Picture segmentation using a recursive region splitting method[J]. *CGIP*，1978，8(3)：313-333.

[54] Tremeau A，Borel N. A region growing and merging algorithm to color segmentation[J]. *Pattern Recognition*，1997，30(7)：1191-1203.

[55] 范立南. 基于多结构元形态边缘检测与不变矩特征模式识别研究[D]. 沈阳：东北大学，2004.

[56] Kass M，Witkin A，Terzopoulos D. Snakes：active contour models[C]. Int.J.Computer

Vision，1987：321-331.

[57] 刘少创，林宗坚. 彩色航空影像分割的OCTOPUS方法[J]. 中国图象图形学报，1997，2(11)：790-794.

[58] 粟海华，王睿斌，庄镇泉. 一种基于判决神经网络的彩色图像分色算法[J]. 模式识别与人工智能，1999，12(3)：345-358.

[59] N Ikonomakis，K N Plataniotis，A N Ventsanopoulos. Grey-scale and colour image segmentation via region growing and region merging[J]. *Canadian Journal of Electrical and Computer Engineering*，1998，28-1(1)：43-47.

[60] 孟章荣. 各种颜色模型选用需求分析[J]. 中国图象图形学报，1996，1(3)：238-241.

[61] Shafarenko L，Petrou M，Kittler J. Histogram-based segmentation in a perceptually uniform color space[J]. *IEEE Trans on Image Processing*，1998，7(9)：1354-1358.

[62] Tominaga S. A color classification method for color image using a uniform color space[C]. Proceeding of the 10th Int. Conf. on Pattern，1990：274-280.

[63] Vlachos T，Constantinides A G. A graph-theoretic approach to color image segmentation and contour classification[C]. The 4th Int. Conf. on Image Processing and its Application，1992：354-362.

[64] Samet H. The quad tree and related hierarchical data structures[J]. *Computer Surveys*，1984，16(2)：187-230.

[65] Gevers T，Kajcovski V K. Image segmentation by directed region subdivision[C]. Proceeding of the 12th IAPR Int. Conf. on Pattern Recognition，1994，963-972.

[66] Loncaric S. A survey of shape analysis techniques[J]. *Pattern Recognition*，1998，31(8)：983-1010.

[67] 席砺莼，闫宏伟. 彩色图像的分割技术[J]. 微机发展，2003，13(4)：46-48.

[68] Healey G E. Segmenting image using normalized color[J]. *IEEE Trans on System，Man，and Cybernetics*，1992，(22)：64-73.

[69] Klinker G J，Shafer S A，Kanada T. A physical approach to color image understanding[J]. *International Journal of Computer Vision*，1990，4(1)：7-38.

[70] Saber E，Tekalp A M，Bozdagi G. Fusion of color and edge information for improved segmentation and edge linking[J]. *Image and Vision Computing*，1997，15：769-780.

[71] Lim Y W，Lee S U. On the color image segmentation algorithm based on the threshold and the fuzzy c-means techniques[J]. *Pattern Recognition*，1990，23(9)：935-952.

[72] Celenk M. A color technique for image segmentation[J]. *Computer Vision*，*Graphics*，*and Image processing*，1990，52(1)：145-170.

[73] Vailaya A，Jain A，Zhang H J. On image classification：city images vs. landscapes[J]. *Pattern Recognition*，1998，31(12)：1921-1935.

[74] Yu H F，Zen H，Wang R L. Color spaces conversion of image between the input and output devices by means of artificial neural network[C]. Proc. 1ICIG，2000：50-53.

[75] Zhang G P. Neural networks for classification：a survey[J]. *IEEE Trans. on Systems*，*Man*，*and Cybernetic-Part C*，2000，30(4)：451-462.

[76] Gonzalez R C，Woods R E. *Digital Image Processing*[M]. 3rd ed. Addison-Wesley，1992.

[77] Serra J. *Image analysis and Mathematical Morphology*[M]. London：Academic Press，1982.

[78] 崔屹. 图像处理与分析数学形态学方法及应用[M]. 北京：科学出版社，2000.

[79] 龚炜，石青云. 数字空间中的数学形态学理论与应用[M]. 北京：科学出版社，1997.

[80] 杨述斌，彭复员. 数学形态学在图像处理中的应用与发展[J]. 武汉化工学院学报，2004，26(1)：70-73.

[81] 吴敏金. 图像形态学[M]. 上海：上海科技文献出版社，1991.

[82] Heath M，Sarkar S，Sanoki T. Comparison of edge detector：a morphology and initial study[J]. *Computer Vision and Image Understanding*，1998，69(1)：38-54.

[83] Parker J R. Special issue on mathematical morphology and nonlinear image processing[J]. *Pattern Recognition*，2000，33(6)：813-819.

[84] Sternberg S R. Grayscale morphology[J]. *Computer Vision*，*Graphics and Image Processing*，1986，35(1)：333-355.

[85] Comer M L，Delp E J. An empirical study of morphological operators in color image enhancement[C]. In Proc the SPIE Conference on Image Processing Algorithms and Techniques III San Jose[A]. California，1992：314-325.

[86] 黄樟灿，陆昊娟. 彩色图像去噪方法探讨[J]. 计算机工程与科学，2001，23(4)：

15-17.

[87] 蒋刚毅，郑义. 关于彩色图像处理的形态学算子研究[J]. 中国图象图形学报，2000，5A (4)：293-299.

[88] Vision J，Sinha D，Douherty E R. Fuzzy mathematical morphology[J]. *Communication and Imagine and Representation*，1992，3(3)：286-302.

[89] 周煦潼，施鹏飞. 模糊数学形态学理论与应用[M]. 上海：上海交通大学出版社，1998.

[90] Koskinen L，A stola J，Neuvo Y. Soft morphological filters[A]. Proc. SPIE Int. Society of Optical Engineering[C]. 1991，15(6)：262-270.

[91] Frank Y. Analysis of the properties of soft-morphological filtering using threshold decomposition[J]. *IEEE Trans. On Signal Proessing*，1995，43(2)：539-544.

[92] 舒昌献，莫玉龙. 基于软形态学的边缘检测[J]. 中国图象图形学报，1999，4A(2)：139-142.

[93] Gasteratos A，Tsalides S. Fuzzy soft mathematical morphology[J]. *Image Signal processing*，1998，145(1)：41-49.

[94] 戴青云，余英林. 一种基于小波与形态学的车牌图像分割方法[J]. 中国图象图形学报，2000，5A(5)：411-415.

[95] 孙杰，张铁群，翟宏琛，唐亦亮. 彩色目标识别中的单通道目标分割方法[J]. 南开大学学报：自然科学版，2002，35(1)：83-87.

[96] Deutsch M，Garcia J，Mandlovic D. Multi-chnnel signal output color pattern recognition by use of a joint transforms correlation[J]. *Applied Ottics*，1996，35(35)：6976-6982.

[97] Zhai Hongchen，Mu guoguang，Sun Jie et al. Color pattern recognition in white – light joint transform correlation[J]. *Applied optics*，1999，38(35)：7238-7244.

[98] Zhai Hongchen，Zhu Xiushan，Mu Guoguang. White light JTC for color pattern recognition[C]. 18th congress of the international commission for optics，San Francisco，USA，1999：213-214.

[99] Munsell A H. *A color notation*[M]. 8th ed. Boston：Munsell Color Company，1939.

[100] Ohta Y，Kanade T，Sakal T. Color information for region segmentation[J]. *Graphics*

Image Process，1980，(3)：241-244.

[101] Huang W C，Wu C H. Adaptive color image processing and recognition for varying backgrounds and illumination conditions[J]. *IEEE Transactions on Industrial Eletronics*，45(2)：351-357.

[102] Moghaddamzadeh A，Bourbakis N. A fuzzy region growing approach for segmentation of color images[J]. *Pattern Recognition*，1997，30(6)：876-881.

[103] Jong Hwan Kim，Hyun Sik Shim，Myung Jin Jung. Cooperative Multi-Agent Robotic Systems[C]：In：TheRobot2Soccer Perspective. Proceedings of the MicroRobot World Cup Soccer Tournament’97. KAIST，Tae-jon，Korea. 1997，3-14.

[104] 佟国峰，高大志，薛定宇，徐心和. 足球机器人视觉系统的设计与开发[C]. 沈阳：东北大学，1998：20-23.

[105] 王文学，赵姝颖，孙萍，等. 多智能足球机器人系统的关键技术[J]. 东北大学学报，2001(4)：192-195.

[106] 何超. 足球机器人视觉图像的快速识别[C]. 上海：华东理工大学出版社，2002.

[107] 赵姝颖. 机器人视觉子系统[J]. 机器人技术与应用，2000(2)：26-28.

[108] 汤顺清. 色度学[M]. 北京：北京理工大学出版社，1990.

[109] 黄晶，赵臣，周明明. 基于快速彩色空间变换的足球机器人目标搜索[J]. 哈尔滨工业大学学报，2003，35(9)：1036-1039.

[110] K S Fu. *Syntactic pattern recognition and application*[M]. Prentice-Hall，Englewood Cliffs，NJ，1982.

[111] P A Devijver，J Kittler. *Pattern recognition：a statistical approach*[M]. Prentice-Hall，Englewood Cliffs，NJ，1982.

[112] E Oja. *Subspace methods of pattern recognition*[M]. Research Studies Press，Letchworth，England，1983.

[113] P A Devijver，J Kittler. *Pattern recognition theory and applications*[M]. Springer Verlag，Berlin-New York-Tokyo，1986.

[114] G Joy，Zhigang Xiang. Reducing false contours in quantized color images[J]. *Computers and Graphics*，1996，20(2)：231-242.

[115] P Heckbert. Color image quantization for frame buffer display[J]. *Computer*

Graphics，1982，16(6)：297-307.

[116] Y Linde，A Buzo，R M Gray. An algorithm for vector quantizes design[J]. *IEEE Transactions on Communications*，1980，28(1)：84- 95.

[117] W H Equitz. A new vector quantization-clustering algorithm[J]. *IEEE Transactions on Acoustics*，*Speech*，*and Signal Processing*，1989，37(10)：1568-1575.

[118] Xiang Zhigang，Joy G. Color image quantization by agglomerative clustering[J]. *IEEE Computer Graphics and Applications*，1994，14(1)：44-48.

[119] M T Orchard，C A Bouman. Color quantization of images[J]. *IEEE Transactions on Signal Processing*，1991，39(12)：2677-2690.

[120] R S Gentile，J P Allebach，E Walowit. Quantization of color images based on uniform color spaces[J]. *Journal of Imaging Techno logy*，1990，16(1)：11- 21.

[121] 赵燕伟，王万良. 基于聚类分析的色彩量化新算法及其应用[J]. 计算机辅助设计与图形学学报，2000，12(5)：340-343.

[122] 刘相滨，邹北骥. 一种新的颜色聚类算法及其图像检索[J]. 计算机工程与应用，2001(13)：126-128.

[123] 于晓波，孙菁，栾新. 基于色度平面的自适应颜色聚类[J]. 青岛海洋大学学报，2002，32(6)：981-986.

[124] 任智斌，隋永新，杨英慧，杨怀江. 在均匀颜色空间中实现彩色图像的颜色量化[J]. 光学，2002，10(4)：340-345.

[125] 童劲松，蔡青. 均匀颜色空间的线性颜色表生成[J]. 计算机辅助工程，1995(4)：48-55.

[126] 刘友明，刘希顺，刘安之，王博亮. 一种基于LUV均匀颜色空间的彩色分割方法[J]. 微型电脑应用，2000，16(12)：27-29.

[127] 艾矫燕，朱学锋. 基于改进算法的ART2网络用于微晶玻璃颜色分类[J]. 华南理工大学学报：自然科学版，2003，31(1)：74-78.

[128] 吴颖谦，胡新荣，杨翠萍，王天珍. 一种新的基于神经网络彩色图像颜色量化方法[J]. 武汉汽车工业大学学报，2000，22(3)：33-36.

[129] Wyszecki G.，Stiles W.S. *Color science*[M]. 2nd ed. New York：John Wiley & Sons，1982.

[130] Judd D B. *Colorimetr y*[M]. National bureau of standards circular 478，1950.

[131] Hunt R W G. *The reproduction of color*[M]. 3rd ed. New York：John Wiley & Sons，1975.

[132] Hunter R S. *The measurement of apperance*[M]. New York：John Wiley & Sons，1975.

[133] Meyer G W. Tutorial on color science[J]. *The Visual Computer*，1986，2：278-290.

[134] Chamberlin. *Colour：its measurement，computation and application*[M]. Heyden，1980.

[135] S Zeki. Colour coding in the cerebral cortex：The reaction of cells in monkey visual cortex to wavelengths and colours[J]. *Neuroscience*，1983，(9)：741-765.

[136] Evan Thompson，Adrian Palacios，Francisco J. Varela. Ways of coloring：Comparative color vision as a case study for cognitive science[J]. *Behavioral and brain sciences*，1992，15(1)：1-74.

[137] D W Hamlyn. *The penguin history of western philosophy*[M]. Penguin，1987.

[138] 刘文耀. 光电图像处理[M]. 北京：电子工业出版社，2002.

[139] 容观澳. 计算机图像处理[M]. 北京：清华大学出版社，2000.

[140] 朱正芳，汤顺青. 色度学及其进展[J]. 物理，1996，25(6)：342-347.

[141] Robert S，Randolph B. *Perception*[M]. 3rd ed. New York，McGraw-Hill，1994.

[142] R L Geogory. *Eye and Brain*.[M]. 3rd ed. London：Weidenfeld and Nicholson，1977.

[143] Elizabeth A，Martin M A. *Concise Medical Dictionary*[M]. 4th ed. London：Oxford University Press，1994.

[144] 刘津，陈奇，俞瑞钊. 计算机颜色科学的发展[J]. 计算机工程，1997，23(2)：47-51.

[145] 田玉敏，梁若莹. 计算机彩色输入输出设备常用颜色空间及其转换[J]. 计算机工程，2002，28(9)：198-200，274.

[146] 田登. 色彩工学[M]. 西安：西安交通大学出版社，1998.

[147] 李维. 颜色特性变量及其交互作用分析[J]. 天津城市建设学院学报，2001，7(3)：204-207.

[148] 刘伟奇，冯睿，周丰昆. 符合人眼视觉特性的颜色亮度模型[J]. 光学学报，

1999，19(10)：1426-1429.

[149] R O Duda，P E Hart. *Pattern classification and scene analysis*[M]. New York：John Wiley and Sons，1973.

[150] Gonzalez R C，Woods R E. 数字图像处理[M]. 阮秋琦，译. 2版. 北京：电子工业出版社，2003.

[151] D C Tseng，C H Chang. Color segmentation using perceptual attributes[C]. Proc. 11th Int.，1992，Conference on Pattern recognition III：Conf C 228-231.

[152] The CMC of the SDC，CMC color-difference formulae[J]. *Color Res. Appl*，1984，(9)：250.

[153] Clarke F J J，McDonald R，Rigg B. Modification to the JPC79 color-difference formula[J]. *J Soc Dyers Color*，1984，(100)：128-132.

[154] Alman D H. Cie TC 1-29 industrial color-difference evaluation[J]. *Color Res. Appl*, 1991，(16)：219-220.

[155] Alman D H. CIE TC 1-29.Industrial color-difference evaluation progress report[J]. *Color Res*, Appl.1993，(18)：137-139.

[156] Jan P Allebach. Processing digital color images：From capture to display[J]. *Physics Today*，1992，(2)：32-39.

[157] Robert M. Boynton. *Human colour vision*[M]. Holt，Rinehart and Winston，1979.

[158] Symon D O. Colour，colour spaces and the human visual system[J]. *Technical report*，*School of Computer Science*，*University of Birmingham*，*England*，1989.

[159] Mark D Fairchild. *Color appearance models*[M]. Massachusetts：Addison Wesley，Reading，1998.

[160] W S Stiles，Gunter W. *Color science concepts and methods*，*quantitative data and formula*[M]. New York：John Wiley and Sons Inc.，2000.

[161] Lin H M，Willson A N. Median filters with adaptive-length[J]. *IEEE CAS1*，1988，35(6)：675-690.

[162] Florencio D A F，Schafer R W. Decision-based median filter using local signal statistics. Proc[J]. *SPIE Symp Visual Comm. Image Processing*，1994，1(23)：268-275.

[163] Hardie R E，Barner K E. Rank conditioned rank selection flters for signal restoration[J].

IEEE IP，1995，3(2)：192-206.

[164] Abreu E，Lightone M，Mitra S K et al. A new efficient approach for the removal of impulse noise from highly corrupted images[J]. *IEEE IP*，1996，5(6)：1012-1025.

[165] Russo F，Ramponi G. A fuzzy filter with adaptive length[J]. *IEEE CAS1*，1988，35(6)：715-720.

[166] Zhang D，Wang Z. Impulse noise detection and removal using fuzzy techniques[J]. *Electronics Letters*，1997，33(5)：378-379.

[167] Kim V，Yaroslavskii L. Rank algorithms for picture processing[J]. *Computer Vision*，*Graphics*，*Image Processing*，1986，35(2)：234-258.

[168] Mitra S K，Yu T H. A new algorithm for the removal of impulse noise from highly corrupted images. Proc[J]. *IEEE CAS*，London，1994：17-20.

[169] Ralph Sucher. A recursive nonlinear filter for removal of impulse noise. Proc. of the IEEE Int. Conf. On Image Processing(ICIP’95)，1995：183-186.

[170] 黄庆华，周荷琴，冯焕清. 一种快速有效的图像脉冲噪声滤除方法[J]. 计算机工程与应用，2002，17(8)：113-115.

[171] 李树涛，王耀南. 图像椒盐噪声的非线性自适应滤除[J]. 中国图象图形学报，2000，5(12)：999-1001.

[172] 孙即祥. 模式识别中的特征提取与计算机视觉不变量[M]. 北京：国防工业出版社，2001.

[173] Christopher M. *Neural networks for pattern recognition*[M]. Oxford：Clarendon Press，1997.

[174] Cybenko G. Approximation by superpositions of a sigmoidal function[J]. *Mathematics of Control*，*Signals*，*and Systems*，1989，(2)：303-314.

[175] Hornik K M，Stinchcombe M，White H. Multi-layer feedforward networks are universal approximators[J]. *Neural Networks*，1989，(2)：359-366.

[176] 徐丽娜. 神经网络控制[M]. 北京：电子工业出版社，2003.

[177] Nielsen R. Theory of the back-propagation neural networks[C]. Proc of the International Joint Conference on Neural Networks，1989：593-611.

[178] 焦李成. 神经网络系统理论[M]. 西安：西安电子科技大学出版社，1990.

[179] 缪绍纲. 数字图像处理[M]. 成都：西南交通大学出版社，2001.

[180] 朱剑英. 智能系统非经典数学方法[M]. 武汉：华中科技大学出版社，2001.

[181] 王耀南，李树涛，毛建旭. 计算机图像处理与识别技术[M]. 北京：高等教育出版社，2001.

[182] 石俊生，杨卫平，白凤翔，张韬，唐伟. 彩色图像分割与颜色特征选择[J]. 云南师范大学学报，1999，19(6)：53-56.

[183] 刘忠伟，章毓晋. 十种基于颜色特征的图像检索算法的比较和分析[J]. 信号处理，2000，16(1)：79-84.

[184] 章毓晋. 基于内容的视觉信息检索[M]. 北京：科学出版社，2003.

[185] Corridoni J M，Del Bimbo A. Film editing reconstruction and semantic analysis[C]. Proc.6ICCAIP，1995，938-943.

[186] 田国会，尹建琴，宁春林，等. 足球机器人视觉子系统及其关键问题[J]. 山东工业大学学报，2002，32(1)：86-91.

[187] 赵姝颖，佟国峰. 微型足球机器人的设计与开发[J]. 机器人技术与应用，1999，1(2)：22-28.

[188] 胡英，徐心和. 色标设计与辨识算法研究[J]. 中国图象图形学报，2002，7(12)：1291-1295.

[189] 阮秋琦. 数字图像处理学[M]. 北京：电子工业出版社，2001.

[190] 陈凤东，洪炳荣，朱莹. 基于HSI颜色空间的多机器人识别研究[J]. 哈尔滨工业大学学报，2004，36(7)：928-930.

[191] 阴炳皓，赵臣，韩晓军. 基于改进的HSI空间模型的目标搜索方法[J]. 河北工业大学学报，2003，32(1)：6-10.

[192] Babenko B，Ming-Hsuan Y，Belongie S. Robust object tracking with online multiple instance learning[J]. *Pattern Analysis and Machine Intelligence*，*IEEE Transactions on*. 2011，33(8)：1619-1632.

[193] T Ojala，M Pietikainen，D Harwood. Performance evaluation of texture measures with classification based on Kullback discrimination of distributions[C]. International Conference on Pattern Recognition(ICPR 1994)，vol.1，pp.582-585.

[194] T Ojala，M Pietikainen，D Harwood. A comparative study of texture measures with

classification based on feature distributions[J]. *Pattern Recognition*，1996，29(1)：51-59.

[195] T Ojala，M Pietikainen，T Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*，2002，24(7)：971-987.

[196] 杨钊. 面向图像分类和识别的视觉特征表达与学习的研究[D]. 广州：华南理工大学，2014.

[197] Papageorgiou C，Poggio T. A trainable system for object detection[C]. IJCV，2000，38(1)：15-33.

[198] 车志富. 基于支持向量机的行人检测[D]. 北京：北京交通大学，2010.

[199] D G Lowe. Distinctive image features from scale invariant keypoints[J].*International Journal of Computer Vision*，2004，60(2)：91-110.

[200] N Dalal，B Triggs. Histograms of oriented gradients for human detection[C]. IEEE Conference on Computer Vision and Pattern Recognition，2005.

[201] 舒锐. 基于梯度方向直方图的人体检测方法的研究[D]. 广州：华南理工大学，2010.

[202] Cortes C，Vapnik V. Support-vector networks[J]. *Machine learning*， 1995，20(3)：273-297.

[203] 宋华军. 基于支持向量机的目标跟踪技术研究[D]. 长春：中国科学院研究生院(长春光学精密机械与物理研究所)，2006.

[204] Lu Zhang，Laurens van der Maaten. Preserving Structure in Model-Free Tracking[J]. *IEEE Transactions on Pattern an Analysis and Machine Intelligence*，2014，36(4)：756-767.

[205] M Andriluka，S Roth，B Schiele. Pictorial structures revisited：people detection and articulated pose estimation[C]. Proc. IEEE Conf. Computer Vision and Pattern Recognition，2009： 1014-1021.

[206] V Ferrari，M Marin-Jimenez，A Zisserman. Pose Search：Retrieving People Using Their Pose[C]. Proc. IEEE Conf. Computer Vision and Pattern Recognition，2009.

[207] Y Yang，D Ramanan. Articulated Pose Estimation Using Flexible Mixtures of Parts[C]. Proc. IEEE Conf. Computer Vision and Pattern Recognition，2011： 1385-1392.

[208] 郑大禹. 基于图的结构化图像特征提取与跟踪[D]. 上海：上海交通大学，2012.

[209] 鲍文霞. 基于结构特征的图像匹配算法及应用[D]. 合肥：安徽大学，2010.

[210] Horowitz E，Sahni S. *Fundamentals of computer algorithms*[M]. Computer Science Press，1978.

[211] 高一凡. 数据结构算法实现及解析[M]. 西安：西安电子科技大学出版社，2002.

[212] 冉玉梅. 基于最小生成树的图像方法研究[D]. 济南：山东师范大学，2010.

[213] D Zhang，S Z Li，D Perez. Real-time face detection using boosting in hierarchical feature spaces[C]. In Int.Conf. on Image Processing，2004.

[214] 李坤. 基于在线学习的视频跟踪算法研究与实现[D]. 长沙：国防科学技术大学，2009.

[215] Feng Tang，S Brennan，Qi Zhao，Hai Tao. Co-tracking using semi-supervised support vector machines[C]. IEEE 11th International Conference on Computer Vision.2007：1-8.

[216] 孙肃肃. 于HOG-PCA和DPM的人脸检测方法的研究和实现[D]. 北京：北京邮电大学，2013.

[217] Zhu M，Wang Y，Chen S，et al. *Sphere-structured support vector machines for multi-class pattern recognition*[M]. Rough Sets，Fuzzy Sets，Data Mining，and Granular omputing. Springer Berlin Heidelberg，2003：589-593.

[218] Crammer K，Dekel O，Singer Y. Online passive aggressive algorithms[J]. *Machine Learning Research*，2006，7(8)：551-585.

[219] 刘建伟，申芳林，罗雄麟. 感知器学习算法研究[J]. 计算机工程，2010，36(7)：190-192.

[220] 肖鹏，段洣毅，赵琪，张可黛. 基于置信图自适应融合的视觉目标跟踪[J]. 无线电工程，2013，43(9)：20-23.